

PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-21314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-235281P.
PR

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1, Page 4168-4169; 11750pp; English.

Claim 1; Page 4168-4169; 11750pp; English

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABY00010-ABY62213) of the
CC specification or its complement. (I) is useful for:
CC

specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer
CC (b) monitoring the progression of prostate cancer in a patient;

(c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC (u) accessing the database for information
CC in a patient; _____

(e) selecting a composition for inhibiting prostate cancer in a patient;

CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient:

CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

AA Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;
 SQ

Query Match 98.8%; Score 2531.6; DB 23; Length 3634;

Best Local Similarity 99.8%; Pred. No. 0;
 Watched 3556; Counted 4; Matches 4; Models 3; Cans 3;

Matchless 2500;	Conservative V;	Matchless 2500;	Conservative V;	Matchless 2500;	Conservative V;
1	1	1	1	1	1
2	2	2	2	2	2
3	3	3	3	3	3
4	4	4	4	4	4
5	5	5	5	5	5
6	6	6	6	6	6
7	7	7	7	7	7
8	8	8	8	8	8
9	9	9	9	9	9
10	10	10	10	10	10
11	11	11	11	11	11
12	12	12	12	12	12
13	13	13	13	13	13
14	14	14	14	14	14
15	15	15	15	15	15
16	16	16	16	16	16
17	17	17	17	17	17
18	18	18	18	18	18
19	19	19	19	19	19
20	20	20	20	20	20
21	21	21	21	21	21
22	22	22	22	22	22
23	23	23	23	23	23
24	24	24	24	24	24
25	25	25	25	25	25
26	26	26	26	26	26
27	27	27	27	27	27
28	28	28	28	28	28
29	29	29	29	29	29
30	30	30	30	30	30
31	31	31	31	31	31
32	32	32	32	32	32
33	33	33	33	33	33
34	34	34	34	34	34
35	35	35	35	35	35
36	36	36	36	36	36
37	37	37	37	37	37
38	38	38	38	38	38
39	39	39	39	39	39
40	40	40	40	40	40
41	41	41	41	41	41
42	42	42	42	42	42
43	43	43	43	43	43
44	44	44	44	44	44
45	45	45	45	45	45
46	46	46	46	46	46
47	47	47	47	47	47
48	48	48	48	48	48
49	49	49	49	49	49
50	50	50	50	50	50
51	51	51	51	51	51
52	52	52	52	52	52
53	53	53	53	53	53
54	54	54	54	54	54
55	55	55	55	55	55
56	56	56	56	56	56
57	57	57	57	57	57
58	58	58	58	58	58
59	59	59	59	59	59
60	60	60	60	60	60
61	61	61	61	61	61
62	62	62	62	62	62
63	63	63	63	63	63
64	64	64	64	64	64
65	65	65	65	65	65
66	66	66	66	66	66
67	67	67	67	67	67
68	68	68	68	68	68
69	69	69	69	69	69
70	70	70	70	70	70
71	71	71	71	71	71
72	72	72	72	72	72
73	73	73	73	73	73
74	74	74	74	74	7

QY 1 GCGCGGCGAGCTCGGAGCTGCTCTTCTGTGGTTCTCTTGCCACCGTCGCT 60

Db 71 GCGCGGCGAGGTCGGAGCTGCTGCTTCTGTTCTCTTGTGGCCGCCGTCGCT 130

[illegible]

01 GICCGCIGCCIGGGGIGLCCAGACAAAGGCAGIAGGGGCCAACAGCACCC
12

xy

Db 131 GTCCGGCTGCCTTGGGCTGCCGAACAGACAAGCGTGGGCCACAGCACCCTCAGAAGCCGA 190

121 CGCAGCTCGACGCAGGGCCGCGCAGGAGGCTGGCGATCGCGTGTCTCGGAGGGCGCCGCGC 180

CEC

Db 191. CGCAGTTCAGCGAGGGCCGGCAGGAGGGTGGCGATCGCGTGTGGAGGGCGCCCGC 230

181 GGGCAGCGGGCGCCAGAGGGGAAAGAGCGGGGGCGCGGTCAAGCCGTGCC 240

Dh 251 GGGCAGGCGGGCCAGAGGGGAAGAAGCGGGGGCGGGGTCAGCCCTGGCC 310

[illegible]

241 GGGCCGGGGGGGATGTCGATGCCCTGACCGGATGCCGCTGCCCGGGGTCTGGGGAGGAGC 300

Db 311 GGGCGGCGGGGATGTCGATGCCCTGACGCGATGCCGCTGCCGGGTCCGGGAGGAGC 370

301 TGAAGCAGCCAAAGGAGATCGAGGACGCCGAGGAAGTACTTCATTGAGGCACCGTCAACCA 360

[illegible]

Db 371 TGAAGCAGGCCAAGGAGATCGAGGACGCCGAGAGTACTCCTTCATGGCCACCCTCACCA 430

361 AGCGCCCAAGAGCAATCCAGTTTGCTGATGACATGCAGGAGTTCACCAATTCCCA 420

[illegible]

DD * J I A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

421 CCAAACTGGCCGAGATCTTTGTCCTCGATCTCACAGTCTCCACTGACAGCTACA 480

Db 491 CCAAACTGGCCGAGATCTTTGTCTCGCTCGATCTCACAGTCTCTCCACTGACAGCTACA 550

Query Match	83.0%;	Score 2126.8;	DB 20;	Length 3616;
Best Local Similarity	97.4%;	Pred. NO. 0;		
Matches 2214;	Conservative	0;	Mismatches 32;	Indels 28;
				Gaps 4;

QY	290	CGSGAGGAGCGTGAAGCAGGCGCCAAAGAGATTCAGAGACGGCCGGAAGATCTCTCATGGC	349
Db	1	CAGGAGGAGCGCTGAAGCAGGCGCCAAAGAGATTCAGAGACGGCCGGAAGATCTCTCATGGC	60
QY	350	CACGCTCACAAAGCGGCCCCAAAGACGAATTCAGATTGGTGATGACATGCAGAGATTTCAC	409
Db	61	CACGCTCACAAAGCGGCCCCAAAGACGAATTCAGATTGGTGATGACATGCAGAGATTTCAC	120
QY	410	CAAAATTCCTCCACCAAACTGSCCGAAGATCTTGTCTCGCTGATCTCACAGTCTCCAC	469
Db	121	CAAAATTCCTCCACCAAACTGSCCGAAGATCTTGTCTCGCTGATCTCACAGTCTCCAC	180
QY	470	TGACAGCTACACTTACGCTGACATCCACACAGATTAAGTACTCTGATGATGATGAGTTTCTCCCG	529
Db	181	TGACAGCTACACTTACGCTGACATCCACACAGATTAAGTACTCTGATGATGATGAGTTTCTCCCG	240
QY	530	AGAGAGAGCAGCAAACTCCAAAGGGCAGACACAAATTTCTGTGTGAAGAACATCAAGCA	589
Db	241	AGAGAGAGCAGCAAACTCCAAAGGGCAGACACAAATTTCTGTGTGAAGAACATCAAGCA	300
QY	590	GGCAGAAATTTGACCGCCGGGAGATTAGATTCGACAGCAAGAATGTCTGTGATTTTC	649
Db	301	GGCAGAAATTTGACCGCCGGGAGATTAGATTCGACAGCAAGAATGTCTGTGATTTTC	360
QY	650	ACTCGAAGAAAGTGCTCAGGGGGAGAGACCCCTGGCGGTGCTAAATTAAGGGGCTGAC	709
Db	361	ACTCGAAGAAAGTGCTCAGGGGGAGAGACCCCTGGCGGTGCTAAATTAAGGGGCTGAC	420
QY	710	ACACATCACAGCCACAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTG	769
Db	421	ACACATCACAGCCACAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGGCTCAGTG	480
QY	770	CCGCTGCTGTGCTTTGTAACATCTACTCAACTCAGAAATGAAGTAGCTCAGACACTGGCTGA	829
Db	481	CCGCTGCTGTGCTTTGTAACATCTACTCAACTCAGAAATGAAGTAGCTCAGACACTGGCTGA	540
QY	830	GGCTGGAGTTGAGATTTTGGCTTGGTGAAGGGCGAGTCGAAGATATGACTTCTGGTGGTGTAT	889
Db	541	GGCTGGAGTTGAGATTTTGGCTTGGTGAAGGGCGAGTCGAAGATATGACTTCTGGTGGTGTAT	600
QY	890	TGACCGCTGTGTGAACATGAGATGGGTGGCGAGGCCCAACATGATCCTGTGATGTGGGGAGA	949
Db	601	TGACCGCTGTGTGAACATGAGATGGGTGGCGAGGCCCAACATGATCCTGTGATGTGGGGAGA	660
QY	950	CTTTAACCCACTGGGTTTATTAAGAAGTATCCAAACGTGTTTAAAGAAATCCGAGGCAATGT	1009
Db	661	CTTTAACCCACTGGGTTTATTAAGAAGTATCCAAACGTGTTTAAAGAAATCCGAGGCAATGT	720
QY	1010	GGAAAGAGACGTGACGTGGGTGTACAGGGCTGTACAGCTCTCCAAACCTGGGAAGCTGTG	1069
Db	721	GGAAAGAGACGTGACGTGGGTGTACAGGGCTGTACAGCTCTCCAAACCTGGGAAGCTGTG	780
QY	1070	TGTTCCGGCCATGAACGTCATGATTTCTGTACCAACAGAAAGTTGATAACTTGTACTG	1129
Db	781	TGTTCCGGCCATGAACGTCATGATTTCTGTACCAACAGAAAGTTGATAACTTGTACTG	840
QY	1130	CTGCGCAGAAATCATTTTGGATGGCCCTGAAGAGGACACAGATGATGATTTGGTGGGGA	1189
Db	841	CTGCGCAGAAATCATTTTGGATGGCCCTGAAGAGGACACAGATGATGATTTGGTGGGGA	900
QY	1190	ACAAGTGTGTGTGTGTGGCTATGTGTGAGTGAAGGCAAGGGCTGCTGTGCTCTCAAGC	1249
Db	901	ACAAGTGTGTGTGTGTGGCTATGTGTGAGTGAAGGCAAGGGCTGCTGTGCTCTCAAGC	960
QY	1250	TCTTGGAGCAATTTGTCTACTATTACCAAAATCGACCCCATCTGTGTCTCTGCAAGGCTGAT	1309
Db	961	TCTTGGAGCAATTTGTCTACTATTACCAAAATCGACCCCATCTGTGTCTCTGCAAGGCTGAT	1020

QY	1310	GGATGGGCTTCAGGGGTGTAAAGCTAATGAAGCATCCGGCAGTGCATGATCTGTAATAC	1369
Db	1021	GGATGGGCTTCAGGGGTGTAAAGCTAATGAAGCATCCGGCAGTGCATGATCTGTAATAC	1080
QY	1370	TTGCACAGGAAATTAATAAGTAGTGACAGGGGAGCATCTTGATGCGATGTAAAAACAGTTG	1429
Db	1081	TTGCACAGGAAATTAATAAGTAGTGACAGGGGAGCATCTTGATGCGATGTAAAAACAGTTG	1140
QY	1430	TATCGTATGCATATATGSGGCCACTCCAAACACAGAAATCGATGTGACACAGCCTCCGACTCC	1489
Db	1141	TATCGTATGCATATATGSGGCCACTCCAAACACAGAAATCGATGTGACACAGCCTCCGACTCC	1200
QY	1490	GGAGCTGACAGTGGGAGCCAGATACGTTCTCAGGTGGACACATGTCATCTGGCCATATGGCAA	1549
Db	1201	GGAGCTGACAGTGGGAGCCAGATACGTTCTCAGGTGGACACATGTCATCTGGCCAGATGGCAA	1260
QY	1550	ACGAGTGTTCCTCTGCGACAGAGGGTGTCTACTCAATTTAGTGTGCTCCACAGATCCAC	1609
Db	1261	ACGAGTGTTCCTCTGCGACAGAGGGTGTCTACTCAATTTAGTGTGCTCCACAGATCCAC	1320
QY	1610	CTTTGTTCTGTCCATCACAAGCCCAACACAGAGGCTTGGCACTGATAGAACTCATATATGC	1669
Db	1321	CTTTGTTCTGTCCATCACAAGCCCAACACAGAGGCTTGGCACTGATAGAACTCATATATGC	1380
QY	1670	ACCCGAGGGGCGATTAACAGAGATGTGTACTTGTTCTCTTAAGAAAATGGATGAATACGT	1729
Db	1381	ACCCGAGGGGCGATTAACAGAGATGTGTACTTGTTCTCTTAAGAAAATGGATGAATACGT	1440
QY	1730	TGCGACCTTGCACTGCGACATATTGTATGGCCACACTACAGAGGTGACAGATGACCAAGC	1789
Db	1441	TGCGACCTTGCACTGCGACATATTGTATGGCCACACTTACAGAGGTGACAGATGACCAAGC	1500
QY	1790	AAATATCTGCGACTCAACAAAAATGGGCCATTCMAACCTAATATTATACAGATACTAATG	1849
Db	1501	AAATATCTGCGACTCAACAAAAATGGGCCATTCMAACCTAATATTATACAGATACTAATG	1560
QY	1850	GACCACTACTACCAAGGACGACAGCTCACCTGAACCAACACACTCTTAAGAAATATTTTAAAG	1909
Db	1561	GACCACTACTACCAAG-----CATTCATATCGGGATGCTTTGGG	1597
QY	1910	ATATACCTTTATTTTCTTCTTACTACCTCTGCTGCTGATTTTTCCTAATATTCATCTT	1969
Db	1598	CCATGCTGCCAGTCCACCTGA--CTTCCCTTGATTTTTCCTAATATTCATCTT	1654
QY	1970	GTTTTTTCATCTCATTTATTCAGATTTGCGACACACAGAGAACTCTTCATGSGCTT	2029
Db	1655	GTTTTTTCATCTCATTTATTCAGATTTGCGACACACAGAGAACTTCTTCATGSGCTT	1714
QY	2030	TAGATGAATAGAGTTGAGGGTCCCTCACTAGTCACTAAGAGAGATTTACTCCC	2089
Db	1715	TAGATGAATAGAGTTGAGGGTCCCTCACTAGTCACTAAGAGAGATTTACTCCC	1774
QY	2090	CACGCCAGAAAGTGATCTTCTCTTTACATTTCTGGGACTTATGCTTAATATAGTA	2149
Db	1775	CACGCCAGAAAGTGATCTTCT--CTTTACCATTTCTGGGACTTATGCTTAATATAGTA	1833
QY	2150	CGTATTAACAGGAAATGCTAAGGTACTCTCTGTGGAACAATCTGCAATGTCTAAATC	2209
Db	1834	CGTATTAACAGGAAATGCTAAGGTACTCTCTGTGGAATAATCTGCAATGTCTAAATC	1893
QY	2210	GCGTTAAAGAGCCATTTCTTAGCTGCTGAATCAAGTGTCTCTTCACTTCTTCAGAGAA	2269
Db	1894	GCGTTAAAGAGCCATTTCTTAGCTGCTGTAATCAAGTGTCTCTTCACTTCTTCAGAGAA	1953
QY	2270	GCAGGAGTGTACTACCCGCGCAGGTAGGTTAATGTGGGTGGGCAATGTTAAATTTCCCT	2329
Db	1954	GCAGGAGTGTACTACCCGCGCAGGTAGGTTAATGTGGGTGGGCAATGTTAAATTTCCCT	2013
QY	2330	TAGAGATTCCAAAGCCCTGTTTCCGCGCTAAAGGAGTATGTCACGTTCCAGAGATGTGAT	2389
Db	2014	TAGAGATTCCAAAGCCCTGTTTCCGCGCTAAAGGAGTATGTCACGTTCCAGAGATGTGAT	2073
QY	2390	AATGAGATGCGTTGTTAAGATCAGAGGCCCACTTGGAATTAATGATATAGCCCTTCTCTC	2449

DB 2074 AATAGATGCTTGTATGATCAGAGCCGACCTGATTTATAGTACCCCTCC 2133
 2450 CACTCCACGACACTGCTATTTTCAGTTTAACTAGACTACACTATTTAGTT 2509
 DB 2134 CACTCCACGACACTGCTATTTTCAGTTTAACTAGACTACACTATTTAGTT 2192
 QY 2510 TAAATTTGCTCTAGATTTATTTCTGTGTCTCAAAAAA 2563
 DB 2193 TAAATTTGCTCTAGATTTATTTCTGTGTCTCAAAAAA 2246
 RESULT 5
 AAS76216/c
 ID AAS76216 standard; cDNA; 5030 BP.
 XX AAS76216;
 AC
 DB 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #13020.
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 PN WC2001.75067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 PI Dmanac RT, Liu C, Tang YT;
 DR WPI: 2001-639362/73.
 PT P-PSDB; ABG12029.
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 PS Claim 1; SEQ ID No 12020; 103bp; English.
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp://wipo.int/pub/published_pcl_sequences](http://wipo.int/pub/published_pcl_sequences).
 XX Sequence 5030 BP; 1432 A; 1137 C; 1178 G; 1283 T; 0 other;

Query Match 37.28; Score 952.4; DB 23; Length 5030;
 Best Local Similarity 73.88; Pred. No. 2.9e-236;
 Matches 1277; Conservative 0; Mismatches 446; Indels 7; Gaps 5;
 QY 149 GGTGGGCGATGGCTGTGCGAGAGGGCGCGCGGAGGGCGCGCGAGGGGGA 208
 DB 4872 GGCAGGGGCTGGAACCTGTAGGGCTCCAGCTCCCGCGGAGCGCGCGCGG 4813
 QY 209 AAGAGGGCGGGCGGGGAGGTAGCGCGCTGAGCGGGCGGGGAGGATGCTGTA 268
 DB 4812 CCCGGGCTCGGGGGCGCGCGCGCGCTCTCAGCGCGCGCGGGAGGTGCTT 4753
 QY 269 CGGATGCGCGTCCCGGGGGTGGGAGAGCTGAACAGCGCCAGAGATGAGAGCG 388
 DB 4752 GGCCATGAAGCGGAGCGAGCCACATCACAGCAGCGGACCGCGGCGGAGGC 4693
 QY 329 C---GAGAAATGCTCTTCATGCGCCACCGTCCACAAAGCGCCCAAGAAATCCAGTT 385
 DB 4692 CCGGTGACGCGCGCGAGCGGACCGTACCGAGCGCGCGCGACATGAAAGATCCAGTT 4633
 QY 386 TGCTGATGACATGACGAGAGTTCACCAAAATCCCAACAAAGTGGCTTTGTC 445
 DB 4632 TGCTGACCAAGAAAGAAATTCACAAAGTCCCAACAAATGGACGCTCTTTGTC 4573
 QY 446 TCGCTGATCTCACAGTCTCCACTGACAGCTACAGTTGAGTCTTACAGATAG 505
 DB 4572 TCGTTCCATTTCTCAGTACATGACTAGCTACAGCTCAGCGGCTTCATATACAGATAG 4513
 QY 506 CTCGTGATATGAGTTTCTCCCGAGAGAGAGCAAACTCCAAAGGCGAGCA 555
 DB 4512 CTCGTGATATGAGATGCGCCAGAGGAGCAAGCAAAAGTCTTAAGGGAAGAGTGA 4453
 QY 566 TTTCTGTGGAAGAAATCAAGCAAGCAAGAAATTTGGAGCGCGGGAATGAGATTGAGCA 635
 DB 4452 CTTCGTGTTAAGAACATCAACAGCAAGCAAGTTGGAGCAAGAAATTTGCTGA 4293
 QY 626 GCAAGACATGCTGCTCTGATTTCACTCAGAAAGCTGCTCAGGGGAGAGCCCTTGGC 685
 DB 4392 ACAAGAAATGCTGCTGATGATGATGCTTTGAGAAAGAGTCAAGAGAAAGCTTTGGC 4333
 QY 686 TGTGCTAAATAGTGGGCTGTACACATCAACAGCCAGACGGGTGTTGATGAGAC 745
 DB 4332 TGGAGCCAAATCGGGGCTTGCACACATCACTGCTCAGAGCTGTGCTTATGGAAC 4273
 QY 746 ACTGTGCTCCCTGGGGGCTCAGTGGCGCTGCTGTGTTACATCTACTCACTGAA 805
 DB 4272 TCTGGGTGCTCTGGGGGCTCAGTGGCGCTGCTGCAACATCTATTCACCTTCA 4213
 QY 806 TGAAGTACCTCAGACACAGCTGAGGCTGAGTGGAGTTCGCTTGGAAAGGCGAGTC 865
 DB 4212 TGAAGTGGCTGCTGCTCTAGCAGAAAGTGGATTTCTGTTTTCCTGGAAGGAGATC 4153
 QY 866 AAGAGATGACTTGTGCTGTATTTAGACCGCTGTGAAATGATGAGGCGAGCGCA 955
 DB 4152 AGAAGATGACTTGTGCTGTATTTAGATGATGATGATGATGAGGCGAGCGCA 4093
 QY 926 CATATCTGATGATGAGGAGACTTAAACCACTGGGTTTATAGAGTATCAAAAGT 955
 DB 4092 CATATCTGATGATGAGGAGACTTAAACCACTGGGTTTATAGAGTATCAAAAGT 4033
 QY 986 GTTAAAGAAATCCAGGCAATTTGTGGAAGAGCGTGAATGTTGTTACAGGCTGATCA 1045
 DB 4032 GTTAAAGAAATCCAGGCAATTTGTGGAAGAGCGTGAATGTTGTTACAGGCTGATCA 3573
 QY 1046 GCTGTCCAAAGCTGGAGAGCTGTGTTCCGCAATGAACGTCATGATTTTACAA 1105
 DB 3972 ACTGTCCAAAGCTGGAGAGCTGTGTTCCGCAATGAACGTCATGATTTTACAA 3913
 QY 1106 ACAGAGTTTATATCTGTAAGTCTGCGGAGATCCATTTGGAGGCGTGAAGAGAC 1165
 DB 3912 ACAGAAATTTATCAACCTCTACTGTTGCGGTGAATCAATTTGATGACTTAAAGAGAC 3853

QY 1166 CACAGATGATGTTGGTGGGAAACAAGTGTGTGTGGCTATGTGATAGGCAA 1225
 DB 3852 AACACAGATGATGTTGGTGGGAAACAAGTGTGTGTGGCTATGTGATAGGCAA 3793
 QY 1226 GGGCGCTGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 1285
 DB 3792 AGGGTGTGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 3733
 QY 1286 CATCTGTGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 1345
 DB 3732 CATCTGTGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 3673
 QY 1346 CCGGCAAG-TCGATGTCTTAAATCTTGAACAGAAATTAAGATGTGTGACGGGAGC 1404
 DB 3672 CCGGCAAGTGTGACATGTCTTAAATCTTGAACAGAAATTAAGATGTGTGACGGGAGC 3613
 QY 1405 ACTTGATGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 1464
 DB 3612 ACTTGATGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 3553
 QY 1465 TCGATGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 1523
 DB 3552 TCGATGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 3493
 QY 1524 GACCATGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 1583
 DB 3492 GACCATGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 3433
 QY 1584 AATTGATGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 1642
 DB 3432 AATTGATGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 3373
 QY 1643 TTTGGCATGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 1702
 DB 3372 TTTGGCATGTGTGGTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 3313
 QY 1703 GCTTCTTAAAGCTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 1761
 DB 3312 GCTTCTTAAAGCTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 3253
 QY 1762 ACCTTAAAGCTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 1821
 DB 3252 ACCTTAAAGCTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 3193
 QY 1822 TCAAACTTAAAGCTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 1871
 DB 3192 TCAAACTTAAAGCTCTTAAAGCTCTTGGAGCAATGTCTACATCTCCGAAATCGAACC 3143

RESULT 6

AAV73925 standard; DNA; 2226 BP.

AAV73925;

04-MAR-1999 (first entry)

Human SAHH DNA #2.

KW S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;
 KW treatment; infection; cancer; autoimmune disease; detection; diagnosis;
 KW gene mapping; antisense; therapy; antagonist; immunosassay; ss.

Homo sapiens.

US5854023-A.

29-DEC-1998.

17-JUL-1997; 97US-0896005.

17-JUL-1997; 97US-0896005.

XX

XX

XX

XX

XX

XX

(INCY-) INCYTE PHARM INC.

Corley NC, Hillman JL, Lal P, Shah P;

WPI; 1999-094906/08.

P-PSDB; AAM90061.

PT Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase -
 PT for production of recombinant enzyme, useful for diagnosis,
 PT treatment and prevention of cancers, infections and autoimmune
 PT diseases

Claim 5; Column 39-42; 40pp; English.

CC This sequence encodes a human S-adenosyl-5-homocysteine hydrolase (SAHH).
 CC The SAHH protein can be used to generate specific antibodies and in drug
 CC screening to identify specific binding agents. Antagonists of the
 CC protein are used to treat or prevent a wide range of viral, bacterial,
 CC fungal, parasitic, protozoal or helminthic infections, many cancers
 CC (leukemia, lymphoma or solid tumours), and many autoimmune diseases
 CC (e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes
 CC mellitus, multiple sclerosis etc). All these conditions may be treated by
 CC expressing antisense sequences, triplex-forming agents or ribozymes
 CC directed against the nucleic acid. The nucleic acid and its fragments can
 CC be used as probes or primers for detecting and quantifying gene
 CC expression, for diagnosis or monitoring of disease, to identify genetic
 CC variations, mutations or polymorphisms. In gene mapping and as antisense
 CC therapeutics. Antibodies are used directly as antagonists, indirectly to
 CC deliver active agents to SAHH-expressing cells, to diagnose and monitor
 CC diseases in standard immunoassays, in competitive drug screens and to
 CC isolate the protein from natural sources.

Sequence 2226 BP; 649 A; 455 C; 435 G; 675 T; 12 other;

Query Match 31.5%; Score 806.2; DB 20; Length 2226;

Best Local Similarity 99.3%; Pred. No. 1.6e-198;

Matches 830; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1728 GTTGCAGCTTGCATCTGCCATCTTTGATGCCACCTTACAGAGCTGACAGATGACCA 1787
 DB 1 GTTGCAGCTTGCATCTGCCATCTTTGATGCCACCTTACAGAGCTGACAGATGACCA 60
 QY 1788 GCAAAATATGTGGAGCAACCAAAAATGGGCAATTCGAACTATTTATACAGATACCA 1847
 DB 61 GCAAAATATGTGGAGCAACCAAAAATGGGCAATTCGAACTATTTATACAGATACCA 120
 QY 1848 TGGACCATATACCAAGGACAGTCCACCTGACACACACTGAAAGAAATATTTTAA 1907
 DB 121 TGGACCATATACCAAGGACAGTCCACCTGACACACACTGAAAGAAATATTTTAA 180
 QY 1908 AGATACTTTATTTCTTACTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1967
 DB 181 AGATACTTTATTTCTTACT 240
 QY 1968 TTGTTTTTCACTCATTAACCAAGTCTGACAGCAACAGGAACTGCTTCAAGGCTC 2027
 DB 241 TTGTTTTTCACTCATTAACCAAGTCTGACAGCAACAGGAACTGCTTCAAGGCTC 300
 QY 2028 TTAGATGAATAGAGTGTAGGCTCCCTGACCTAGTACTTAAAGAGATTTTACTCC 2087
 DB 301 TTAGATGAATAGAGTGTAGGCTCCCTGACCTAGTACTTAAAGAGATTTTACTCC 360
 QY 2088 CCCAGCCAGAGAAAGTATCTCTCTTAAACATTTCTGGGAGCTTAACTTAAATAGG 2147
 DB 361 CCCAGCCAGAGAAAGTATCTCTCTTAAACATTTCTGGGAGCTTAACTTAAATAGG 419
 QY 2148 TACCTTATTAAGAGAAATCTAAGTACCTCTCTCTGGAACAATCGCAATGCTTAA 2207
 DB 420 TACCTTATTAAGAGAAATCTAAGTACCTCTCTCTGGAACAATCGCAATGCTTAA 479
 QY 2208 TCGCCTTAAAGAGCCATTTCTAGCTGCTGAATCAGTCTCTTCTCACTTCTTCAAG 2267
 DB 480 TCGCCTTAAAGAGCCATTTCTAGCTGCTGAATCAGTCTCTTCTCACTTCTTCAAG 539

QY	23268	AACGAGGATGAGTACTACCCGGCAGGTAGTATGAGGGTGGCATGTGTAATTTCC	2327
Db	540	AACGAGGATGAGTACTACCCGGCAGGTAGTATGAGGGTGGCATGTGTAATTTCC	599
QY	23238	CTTAGAAGTTCACAGCCCTGTGTTCTCTGCGTAAAGGTGTATGTCCAGATGATGT	2387
Db	600	CTTAGAAGTTCACAGCCCTGTGTTCTCTGCGTAAAGGTGTATGTCCAGATGATGT	659
QY	23388	ATTATGAGCATGGCTGTTAAGATTCAGAGAGGCCAATTGGATTTATAGATAGCCCTCC	2447
Db	660	ATTATGAGCATGGCTGTTAAGATTCAGAGAGGCCAATTGGATTTATAGATAGCCCTCC	719
QY	24448	TCCACTCCACACAGACTTGCTCATTTTTCGAGTTTTTAACAGACTACACTGTATTTGAG	2507
Db	720	TCCACTCCACACAGACTTGCTCATTTTTCGAGTTTTTAACAGACTACACTGTATTTGAG	778
QY	2508	TTTAAATTTTCTCTCAGAGATTAAATTCGTGTCCAAAAAAAAAAAAAAAAAAAAA	2563
Db	779	TTTAAATTTTCTCTCAGAGATTAAATTCGTGTGTCCAAAAAAAAAAAAAAAAAAAAA	834

SQ Sequence 721 BP; 157 A; 195 C; 243 G; 119 T; 7 other;
XX
CC
CC gastroenterostinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AA16506 to AA16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX

	Query Match	Similarity	27.4%	Score	702.4	DB	21	Length	721
	Best Local	Similarity	98.0%	Pred. No.	7.5e-172				
	Matches	703	Conservative	6	Mismatches	8	Indels	0	Gaps
QY	1	GGCGCGGCGAGGTGCGAGCTCGAGCTGCTGTTGTTGTTCTGTGGCCACCGTCGT	60						
Db	4	GGCGCGGCGAGGTGCGAGCTCGAGCTGCTGTTGTTGTTCTGTGGCCCGCCGTCGT	63						
QY	61	GTCCGGCTGCTTGGGCTGCCGACAGACAAAGCGTGGGCCACAGACACTCAGAAAGCGA	120						
Db	64	GTCCGGCTGCTTGGGCTGCCGACAGACAAAGCGTGGGCCACAGACACTCAGAAAGCGA	123						
QY	121	CGCACTGACAGCGAAGGCGCGGACAGAGAGGTGGCGATGCTGTTCGAGAGGCGCCGCCGC	180						

RESULT 7
AAFI6004
ID AAFI6004 standard; cDNA; 721 BP
XX

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:439.

KM Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KM neuroprotective; cyrostatic; cardioactive; immunomodulatory; muscular;
 KM vulnereary; gastrointestinal; nephrotoxic; antinefactive; gynecological
 KM antibacterial; gene therapy; neural; immune; reproductive; renal;
 KM gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KM wound; infectious disease; ss.

Homo sapiens.

PN W0200055174-A1

PD 21-SEP-2000

PF 08-MAR-2000; 2000WO-US05988

PR 12-MAR-1999; 99US-0124270

PA (HUMA-) HUMAN GENOME SCI INC.

XX XX

2000

DR P-PSDB; AAB56801.

Prostate cancer a

PT disorders such as prostate cancer -

PS Claim 1; Page 917-918; 2338pp; English

CC AAF15566 or AAF15505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardiactive, immunomodulatory, muscular, vulnary, gastrointestinal,
CC neurotropic, antinefactive, gynaecological and antibacterial activities
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer,
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to
CC disorders such as neural, immune, muscular, reproductive,

Query Match	27.4%	Score 703.4	DB 21	Length 721	.
Best Local Similarity	98.0%	Pred. No. 7.5e-172			
Matches 703, Conservative	6	Mismatches 8		Indels 0	Gaps 0

SQ Sequence 721 BP; 157 A; 195 C; 243 G; 119 T; 7 other;
XX
CC
CC gastro-intestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AA16506 to AA16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX

181 GGCGAGCGGGCGGGCGCCAGAGGGGGAAGAGCGGGGGCGGGCGGTCAAGCCGTGGCC 240

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Db	304	TGAAGCAGGCCAAGGAGATCGAGACGCCGAAAGTCTCCTTCATGCGCCACCGCTACCA	365
QY	361	AGGGCCCAAGAAACCAATCCAGTTGCTGATGCATGCAGAGAGTTGACCAAAATTCCCA	420
Db	364	AGGGCCCAAGAAACCAATCCAGTTGCTGATGCATGCAGAGAGTTGACCAAAATTCCCA	423

QY 421 CCAAAZ

Db 424 CCAAAACT

QY 481 GTTCA

Db 484 GTTCAGCTGCATCCTACACAG

QY 541 AAACCACTCCAAGGGCAGCA

Db 544 AAACCACTCCAAAGGGCAGCAG

QY 601 GACGCCGGAGATTGAGATTGCAGAGCAAGACATGTCCTCTGAT

Db 604 GACGCCGGGAGATTG

QY 661 GTGCTCAGGGGGAGA

Db 664 GTGCTCAGGGGARAAGCC

ABL03693

XX	ABL03693;
AC	
XX	26-MAR-2002 (first entry)
DT	
XX	
XX	Drosophila melanogaster expressed polynucleotide seq ID NO 5561.
DE	
XX	
XX	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical; gene; ss.
XX	

20

OS Drosophila melanogaster.
 XX MO200171042-A2.
 PN 27-SEP-2001.
 PD 23-MAR-2001; 2001MO-0509231.
 PF 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI MPI; 2001-656860/75.
 DR P-PSDB; ABB59590.
 XX
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 PS Claim 1; SEQ ID NO 5561; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 1566 BP; 360 A; 438 C; 481 G; 287 T; 0 other;
 SO
 Query Match 27.28; Score 696.2; DB 23; Length 1566;
 Best Local Similarity 70.58; Pred. No. 4.6e-170;
 Matches 945; Conservative 0; Mismatches 393; Indels 3; Gaps 1;
 QY 509 TGAATGATGAGTTCCTCCCGAGAGAGACGAAACCAATCCAGAGGAGAGCAATTT 568
 DB 225 TGGCGAGAGAGTGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 284
 QY 569 CTGTGTGAAGACAT---CAAGCAGCAGAAATTTGGAGCGCGGAGATTGAGATTGCGA 625
 DB 285 CTGTGTGAAGACATAGGCGCCAGCAGATGATTTGGCGCCAGGAGATGAGATTGCGGA 344
 QY 626 GCAAGACATGTCTGTGATTTCACTCAGAGAAAGTGTCTAGGGGAGAGAGCCCTTGGC 685
 DB 345 GCAAGAGATGCGCGGATCATATAGCGCTAGAGAGCGGCGGAGAGAGAGAGAGAG 404
 QY 686 TGTGTCAAAATATATGGCTTACACACATCAGACCCAGAGAGCGGTGTGATTAGAC 745
 DB 405 GGAATGCAAGATGCTGTGCTTACACACATCAACGCCAGCTGCGTCTCATTCAGAAC 464
 QY 746 ACTGTGCTGCTGGGGGCTCAGTGGCGGTGTGCTCTTAAACATTAATCAATCAGAGAA 805
 DB 465 GCTGTGGAAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 524
 QY 806 TGAAGTACGCTGACACATGCTGAGGCTGAGATTGCAAGTTCGCTTGAAGGGGAGTC 865
 DB 525 CGAAGTGGCTGCTGCTTGGCGAATCCGGAATCCGATCTTGGCTGGCGGCGAGAC 584
 QY 866 AGAAGATGACTTCTGTGTGTATATGACCGCTGTGTGAACATGAGTGGGAGAGCCAA 925
 DB 585 GGAAGAGACTTCTGTGTGTGTATGATGATGATGATGATGATGATGATGATGATGAT 644
 QY 926 CATGATCTGATGATGAGGAGAGAGACTTAACCACTGGGTATTAAAGATCAAAAGT 985
 DB 645 CATGATCTGATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 704

QY 986 GTTTAAGAGATCCGAGCATTTGTGGAAGAGAGCGTGTGCTTACAGAGCTGTATCA 1045
 DB 705 GTTCAGATGTTGTTCAAGGAATCTGTGAGAGAGCGCTTACCGAGAGCGTCTATCA 764
 QY 1046 GCTGTCAAGAGCTGGGAAGCTGTGTCCGGCCATGAAACGATGATGATGATGATGAT 1105
 DB 765 GCTGTCAAGAGCTGGGAAGCTGTGTCCGGCCATGAAACGATGATGATGATGATGATGAT 824
 QY 1106 ACAGAGATTTGATTAATCTGTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1165
 DB 825 GACCAATTTGATTAATCTGTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 884
 QY 1166 CACAGATGTGATGTTGTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1225
 DB 885 CACAGATGTGATGTTGTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 944
 QY 1226 GGGCTGCGTCTGCTGCTCAAGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1285
 DB 945 GGGATGTGCTGAGGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1004
 QY 1286 CATCTGTGCTGTGAGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1345
 DB 1005 CATTTGGCGCTGAGAGCGCAGATGATGATGATGATGATGATGATGATGATGATGAT 1064
 QY 1346 CCGGCAAGTCAATGTCTGATTAATCTTGCACAGAGAGAGAGAGAGAGAGAGAGAGAG 1405
 DB 1065 CCGGCAAGTCAATGTCTGATTAATCTTGCACAGAGAGAGAGAGAGAGAGAGAGAGAG 1124
 QY 1406 CTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1465
 DB 1125 CATGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1184
 QY 1466 CGATGTGACAG 1525
 DB 1185 CGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1244
 QY 1526 CCATGTCTGTGAG 1585
 DB 1245 TCAATATTATGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1304
 QY 1586 TTTGAGCTGTGACAG 1645
 DB 1305 TCTGAGCTGTGACAG 1364
 QY 1646 GGAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1705
 DB 1365 GGAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1424
 QY 1706 TCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1765
 DB 1425 GCGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1484
 QY 1766 TACAGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1825
 DB 1485 GACGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1544
 QY 1826 ACCTAATTATTAACAGATCTA 1846
 DB 1545 GCCCAACTAGTACCGCTACTA 1565

RESULT 9
 ABL29787
 ID ABL29787 standard; DNA; 1776 BP.
 XX
 AC ABL29787;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 40834.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical; gene: ds.
 XX Drosophila melanogaster.
 XX WO2001/1042-A2.
 XX 27-SEP-2001.
 PD 23-MAR-2001; 2001WO-US09231.
 PF 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PMD, Myers EM;
 DR WPI: 2001-656860/75.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 40834; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB161175) and the encoded proteins
 CC (AB057737-AB072072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pt_sequences.
 XX
 XX Sequence 1776 BP; 437 A; 432 C; 479 G; 428 T; 0 other;
 Query Match 25.3%; Score 649.2; DB 23; Length 1776;
 Best Local Similarity 67.0%; Pred. No. 7.8e-158;
 Matches 937; Conservative 0; Mismatches 458; Indels 3; Gaps 1;

Db 695 CGAGAGAGATTCTGCTGCTGCTGACAGAGGCCACTTACTCCGACGCTGCGACCGCA 754
 Qy 925 ACATGATCCTGGATGATGAGGGGAGACTTAACCCACTGGGTTATAGAGATATCCAAACG 984
 Db 755 ACCTATCTCTGACGACGCGGGGATGCGACGACCTTATGCTCAAGATACCCGCACT 814
 Qy 985 TGTTAAGAGATCCGAGGCAATTTGGAGAGAGCGGTACTGTGTTCACAGGCTGTATC 1044
 Db 815 ACTTCAGAGCCATTCGGGCGATCTGAGAGAAAGTGGACCGGGGTGCACCGCTGTACA 874
 Qy 1045 AGCTCTCAAACCTGGGAGGCTCTGTGTTCCGGCCATGAACTCAATGATTTCTTTACA 1104
 Db 875 TGTCTCAAAAGGGCGAAACTACTGTTCCGGCCATGAACTCAATGATTTCTTTACA 934
 Qy 1105 AACAGAGTTGATTAATCTGTAATGCTGCGAGAAATCCATTTGGATGAGGCTGAAGAGA 1164
 Db 935 AGACAGATTGATTAATCTGTAATGCTGCGAGAAATCCATTTGGATGAGGCTGAAGAGA 994
 Qy 1165 CCACAGATGATGATGTTGGTGGGAAACAGTGGTGTGTGCTATGTAGAGTACGCA 1224
 Db 995 CCACGATATTAATGTTGGCGGAAAGAGTGGTATCTGTGGGTACGATGATGGGAA 1054
 Qy 1225 AGGCTGCTGCTGCTGCTCAAAAGCTTTGGAGCAATTTGCTATTAACCGAATTCGACC 1284
 Db 1055 AGGGCTGTGCCAGATCCCTGAAAGGCCCAAGATGATGTTTATGTAAGCAAGTGAATC 1114
 Qy 1285 CCATCTGCTCTGTCAGGCGCTGATGATGGTTCAGGCTGATTAAGCTTAATGAATCA 1344
 Db 1115 CCATATGCTCTTACAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1174
 Qy 1345 TCCGCAAGTCATGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1404
 Db 1175 TCAGGAGCGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1234
 Qy 1405 ACTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1464
 Db 1235 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1294
 Qy 1465 TCGATGTGACAGCCCTCCGACATCCGAGCTGACGTGGAGCGAGTACGTTCTGAGGTGG 1524
 Db 1295 TGTATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1354
 Qy 1525 ACCATGTATCTGCGGCGAGATGCGAAGAGTGTCTCTGCGGAGAGAGGTCTGATCTCA 1584
 Db 1355 ATCCATCATGAGTGGCGCGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGAT 1414
 Qy 1585 ATTGAGCTGTCACAGTCCGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1644
 Db 1415 ATTGCTCTGCTTCCACATTTCTCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1474
 Qy 1645 TGGCACTGATAGACTCTATTAATGCAACCGAGGGGATACAAAGAGATGTGTACTTGC 1704
 Db 1475 TGGCCCTGATTAACCT 1531
 Qy 1705 TTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1764
 Db 1532 TGGCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1591
 Qy 1765 TTACAGAGCTGACAGTACCAAGAAATATCTGGGACATCAAAAGAGGCGATTC 1824
 Db 1592 TCACGAGCTCACGAGTACAGCTCCAAAGTTATGGGCTTAACAAAGCGGCGCTTTTA 1651
 Qy 1825 AACCTAATTATTAAGAT 1842
 Db 1652 AAGCAATTACTACAGGT 1669
 RESULT 10
 ABL79045
 ID ABL79045 standard; cDNA; 553 BP.
 XX ABL79045;
 XX

QY 1267 ACATTACCCAAATCGACCCATGCTGCTGAGCGCTGATGATGGTTCAGGGNG 1326
 DB 3726 AGTTTACCGAGAGTGGATCCCATATGCTCTACAGCTGCATGATGATGGATGGG 3785
 QY 1327 TAAAGTAATGATGATCGGCAAGTCGATGCTGTAATTAATTCAGCAAAATAGA 1386
 DB 3786 TAGGCTCAACGAGGATCAGACAGCGTGTGATGGTGGTACGCAACTGGAAACAAA 3845
 QY 1387 AGTGTAGTACACGGGAGCCTTGATGCGATGAAACAGTTGATGATGATGATG 1446
 DB 3846 AGTGTATTCACGAGGATGATGATGATGATGATGATGATGATGATGATGATG 3905
 QY 1447 GCCACATCCAAACAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1471
 DB 3906 GACATTCCTGCTGCGAGATGATGATGATGATGATGATGATGATGATGATGATG 3965
 QY 1472 -----GACACGCTCCGCACTCCGAGCTCGAGCTGAGTGG 1503
 DB 3966 TATATCAAAAGTTTAAACCACTTAAAGATGGCTGATACCCCGAGCTAAGTGG 4025
 QY 1504 AGCGATAGTGTTCAGGTGAGCATGTCATCTGCGCAGATGAGCAAGTATGCTCC 1563
 DB 4026 AGCGTGTGCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4085
 QY 1564 TGGCAGAGGTCCTACTGATGATGATGATGATGATGATGATGATGATGATGATG 1623
 DB 4086 TGGCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4145
 QY 1624 TCACAGCCACACACAGGCTTGGCACTGATGATGATGATGATGATGATGATG 1683
 DB 4146 TGGCCTCATACACAGGCTTGGCACTGATGATGATGATGATGATGATGATG 4202
 QY 1684 ACAAGCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1731
 DB 4203 ATGATGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4250
 RESULT 14
 AA280766/c
 ID AA280766 standard; cDNA; 636 BP.
 AC AA280766;
 XX
 XX
 DT 07-APR-2000 (first entry)
 XX
 XX
 DE Human colon cancer cell line SW480 cDNA clone SEQ ID NO:850.
 XX
 XX
 KW Human; gene expression product; diagnosis; tumour; colon cancer;
 KW colorectal adenocarcinoma; cell line SW480; cell proliferation;
 KW cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
 KW hyperplasia; ds.
 XX
 XX
 OS Homo sapiens.
 XX
 PN MO9964576-A2.
 XX
 PD 16-DEC-1999.
 XX
 PF 09-JUN-1999; 99WO-IB01062.
 XX
 PR 10-JUN-1998; 98US-0088801.
 XX
 PA (FARB) BAYER CORP.
 XX
 PI Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;
 PI Carroll E, Catino TJ, Derfl A, Ford DM, Lewis ME, Monahan JE;
 PI Schlegel R;
 XX
 WI: 2000-087220/07.
 XX
 PT Novel nucleic acids, used to develop products for the diagnosis and
 treatment of disorders involving unwanted cell proliferation,

PT particularly cancers, especially colon cancer
 XX
 XX Claim 15; Page 469; 469pp; English.
 XX
 CC AA279917 to AA280766 represent double stranded cDNA clones isolated from
 CC the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
 CC cDNA clones can be used to generate antisense oligonucleotides which
 CC can be used for antisense therapy. Methods and products from the present
 CC invention can be used for identifying and/or classifying cancerous cells
 CC present in a human tumour, particularly in solid tumours, e.g.
 CC carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
 CC can be used for developing agents for the diagnosis and treatment of
 CC disorders involving unwanted cell proliferation, such as neoplasia,
 CC dysplasia or hyperplasia.
 XX
 XX Sequence 636 BP; 156 A; 166 C; 137 G; 147 T; 30 other;
 SQ
 Query Match 15.9%; Score 406.4; DB 21; Length 636;
 Best Local Similarity 99.5%; Pred. No. 4.5e-95;
 Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 707 TACACATACACACCCACAGACAGC-GGTGTTGATGAGACACTGTGCTGGGGCTC 765
 DB 420 TACACATACACACCCACAGACAGC-GGTGTTGATGAGACACTGTGCTGGGGCTC 361
 QY 766 AGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
 DB 360 AGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 301
 QY 826 CTGAGGCTGAGTTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
 DB 300 CTGAGGCTGAGTTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
 QY 886 GTATTATCCGCTGTGGAACATGATGATGATGATGATGATGATGATGATGATGATG 945
 DB 240 GTATTATCCGCTGTGGAACATGATGATGATGATGATGATGATGATGATGATGATG 181
 QY 946 GAGACTTAACCCACTGGTGTATAGAGATGATGATGATGATGATGATGATGATGATG 1005
 DB 180 GAGACTTAACCCACTGGTGTATAGAGATGATGATGATGATGATGATGATGATGATG 121
 QY 1006 TGTGTGAGAGAGAGCGTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065
 DB 120 TGTGTGAGAGAGAGCGTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 61
 QY 1066 TGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
 DB 60 TGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1
 RESULT 15
 AAC77823
 ID AAC77823 standard; cDNA; 2200 BP.
 AC AAC77823;
 XX
 XX
 DT 08-FEB-2001 (first entry)
 XX
 XX
 DE Human cancer associated gene sequence SEQ ID NO:217.
 XX
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;
 KW antidiabetic; antihistaminic; antineoplastic; antitubercular; antiviral;
 KW antiinflammatory; antihypertensive; antiallergic; antibacterial; cardiotonic;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotonic; antiproliferative; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.
 XX
 OS Homo sapiens.
 XX
 PT

OM nucleic - nucleic search, using sw model

(w/linac alignments)
11422.740 Million cell updates/sec

Perfect Score: 2503
Sequence: 1 ggcgcgycaggtcgcgagct.....aaaaaaaaaaaaaaaaaaaaa 2503

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

Database

```

GenEmbl: *
1: gb.ba: *
2: gb.htg: *
3: gb.in: *
4: gb.om: *
5: gb.ov: *
6: gb.pat: *
7: gb.ph: *
8: gb.pl: *
9: gb.pr: *
10: gb.ro: *
11: gb.sta: *
12: gb.sy: *
13: gb.un: *
14: gb.vi: *
15: em.ba: *
16: em.fun: *
17: em.hum: *
18: em.in: *
19: em.mu: *
20: em.om: *
21: em.or: *
22: em.ov: *
23: em.pat: *
24: em.ph: *
25: em.pl: *
26: em.ro: *
27: em.sta: *
28: em.un: *
29: em.vi: *
30: em.htg.hum: *
31: em.htg.lnv: *
32: em.htg.other: *
33: em.htg.mus: *
34: em.htg.pln: *
35: em.htg.rod: *
36: em.htg.mam: *
37: em.htg.vrt: *
38: em.sy: *
39: em.htgo.hum: *
40: em.htgo.mus: *
41: em.htgo.other: *

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result

No.	Score	Match Length	DB	ID	Description
1	2563	100.0	2563	6	AX029176 Sequence
2	2563	100.0	2567	6	AF315818 Homo sapi
3	2222	86.7	2552	7	BC016942 Homo sapi
4	2177	84.9	2526	9	BC007576 Homo sapi
5	2127	83.0	2583	9	BC010681 Homo sapi
6	1866	72.4	2258	9	HS082761 Homo sapien
7	1856	72.4	2510	9	HM8000298 Homo sapi
8	1080	42.1	1369	9	AF035319 Homo sapi
9	1035	40.4	1323	9	BC003631 Homo sapi
10	553	21.6	2226	6	AF068569 Sequence
11	553	21.6	2226	6	AX059339 Sequence
12	446	17.4	81378	9	AL772411 Human DNA
13	446	17.4	183584	30	AL160171 Human DNA
14	330	12.9	376	6	AX408384 Sequence
15	227	9.7	292	6	G3161 DRES45 Huma
16	222	8.7	494	11	G3161 DRES45 Huma
17	168	6.6	384	6	AX407699 Sequence
18	152	5.9	1772	10	BC018218 Mus muscu
19	109	4.3	209300	10	AC090750 Mus muscu
20	109	4.3	211438	2	AC122902 Mus muscu
21	81	3.2	1998	9	AF090905 Homo sapi
22	54	2.1	85472	2	AC055888 Rattus no
23	29	1.1	1987	9	AK029572 Homo sapi
24	29	1.1	2301	10	AF476964 Rattus no
25	29	1.1	200801	2	AL773601 Danio rer
26	29	1.1	251429	2	AC114565 Mus muscu
27	29	1.1	257119	2	AC122822 Mus muscu
28	28	1.1	1398	8	AF239829 Petroseli
29	28	1.1	187013	2	AC116920 Homo sapi
30	28	1.1	205804	2	AC125450 Dictyoste
31	27	1.1	1631	9	AF226052 Mus muscu
32	27	1.1	1654	9	AF226052 Homo sapi
33	27	1.1	2294	8	AK000439 Homo sapi
34	27	1.1	69562	8	AY070057 Arabidops
35	27	1.1	74842	8	AC101403 Mus muscu
36	27	1.1	100310	2	AB016877 Arabidops
37	27	1.1	125166	9	AL056652 Arabidops
38	27	1.1	198429	8	AL358613 Human DNA
39	26	1.0	207	6	AT161514 Arabidops
40	26	1.0	787	6	AX186770 Sequence
41	26	1.0	919	9	124089 Sequence 12
42	26	1.0	945	8	AK000027 Homo sapi
43	26	1.0	949	9	Z47980 L. esculentu
44	26	1.0	1165	6	BC029402 Homo sapi
45	26	1.0	1204	9	AX299437 Homo sapi
					BC014356 Homo sapi

ALIGNMENTS

RESULT 1				
AX029176	AX029176	2563 bp	DNA	linear
LOCUS	AX029176			PAT 16-SEP-2001
DEFINITION	Sequence 1 from Patent WO9814562.			
ACCESSION	AX029176			
VERSION	AX029176.1	GI:10190060		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 2563)			
AUTHORS	Hart, D. N.			
TITLE	Enzyme having s-adenosyl-L-homocysteine hydrolase (ahcy) type activity			

JOURNAL Patent: WO/9814562-A 1 09-APR-1998;
HART DEREK NIGEL JOHN (NZ)

FEATURES Location/Qualifiers
Source 1. 2563
/organism="Homo sapiens"
/db_xref="taxon:9606"

CDS
/note="Open reading frame extends without a stop codon for the full 5' nucleotide sequence. The initiation codon has yet to be identified."
/codon_start=1
/protein_id="CAC09285.1"
/db_xref="GI:10190061"

/translation="RGVYARSCCFWFGSGHRCRPAALGCRITKANATAPOKFTOLDAGARVDRVSGAARAGRAPDEGEGGSGAAGAGGMSMPDAPLPGVEELKQAAKEIDAEKYSMATVYKAPKIOIPADMOEFTKPTKGRSLRSLISQSDSYLSAASTYDSDESVPREKQNTSKSGCNKIKOAEGRREIEAOMDSILSRKRAOGERPLAGKIVGCHITTAOTAVLIEICLALGACRNASANITYSTOREVAALAEAGVAIPANKGESEDFWMCIDRCVNMDSQANMLDGDGDLTHVYKRYNPKYKRIAGIYEEVSQVGVGEGVAGCCALALALAIYITIEIDICALQACMDGRVVKLEVIHQVDVITTCGNKNVYRREHLDRMKNSCIVCMGHSNTEIDVTSRTPELTERYRSQVDHYMPDGRVYVLAAGRLMLSCSTPVTLSITATLALALIELYNAPRGRYKODVYLLPKMDEVASLHLPSFDAHLTELDQAKYGLKKNQPFKNYRY"

BASE COUNT 646 a 604 c 677 g 636 t
ORIGIN

Query Match 100.0%; Score 2563; DB 6; Length 2563;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGCGGGGAGAGTGGAGTGGAGTGGTCTTGTCTGTGTGGCCACCGTGGT 60
DB 1 GGGCGGGGAGAGTGGAGTGGAGTGGTCTTGTCTGTGTGGCCACCGTGGT 60
OY 61 GTCCGGCTCTTGGCTCCGACAGACAGAGCGTGGCCACAGCACTCAGAACGGA 120
DB 61 GTCCGGCTCTTGGCTCCGACAGACAGAGCGTGGCCACAGCACTCAGAACGGA 120
OY 121 GCGAGTCCAGCAGAGGCGCGGAGAGAGGTCGCGTGTGCGAGAGGCGCGCGC 180
DB 121 GCGAGTCCAGCAGAGGCGCGGAGAGAGGTCGCGTGTGCGAGAGGCGCGCGC 180
OY 181 GGGCAGGCGGGGGGCGGAGAGAGGAGGAGGCGGGGGGTCAGCCGCTGGCC 240
DB 181 GGGCAGGCGGGGGGCGGAGAGAGGAGGAGGCGGGGGGTCAGCCGCTGGCC 240
OY 241 GGGCGGGGAGATGTGATGCTGACGCGATGCGGTCGCCGGGGTGGGAGAGC 300
DB 241 GGGCGGGGAGATGTGATGCTGACGCGATGCGGTCGCCGGGGTGGGAGAGC 300
OY 301 TGAAGCAGGCGCAAGAGATGAGAGCGCGAGAAATGTCCTTCATGCGCACTCACA 360
DB 301 TGAAGCAGGCGCAAGAGATGAGAGCGCGAGAAATGTCCTTCATGCGCACTCACA 360
OY 361 AGGCGCGGCAAGAGCAATGCTGATGATGATGAGAGAGTTCACCAAAATTCGCCA 420
DB 361 AGGCGCGGCAAGAGCAATGCTGATGATGATGAGAGAGTTCACCAAAATTCGCCA 420
OY 421 CCAAAAGTGGCGAAGATTTGTCTGCTGATGATGATGATGATGATGATGATGATGAT 480
DB 421 CCAAAAGTGGCGAAGATTTGTCTGCTGATGATGATGATGATGATGATGATGATGAT 480
OY 481 GTTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GTTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
OY 541 AAACCACTCCAAAGGCGAGCAATTTCTGTGAAGAATCAATCAAGCAGGAGAAATTTG 600
DB 541 AAACCACTCCAAAGGCGAGCAATTTCTGTGAAGAATCAATCAAGCAGGAGAAATTTG 600
OY 601 GAGCGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 GAGCGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660

DB 601 GAGCGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
OY 661 GTGCTCAGGAGGAGAGACCCCTTGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GTGCTCAGGAGGAGAGACCCCTTGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
OY 721 CCCACAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 721 CCCACAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
OY 781 CTTGTACATCTACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 840
DB 781 CTTGTACATCTACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 840
OY 841 CAGTGTTCCTTGAAGAGGCGAGTCAAGAGATGATGATGATGATGATGATGATGATGATG 900
DB 841 CAGTGTTCCTTGAAGAGGCGAGTCAAGAGATGATGATGATGATGATGATGATGATGATG 900
OY 901 TGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 TGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
OY 961 GGGTTATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 GGGTTATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
OY 1021 TGACTGTGTTACAGAGCTGTATCAGCTCTCCAAAGCTGGAGAGCTGTGTTCGGGCA 1080
DB 1021 TGACTGTGTTACAGAGCTGTATCAGCTCTCCAAAGCTGGAGAGCTGTGTTCGGGCA 1080
OY 1081 TGAAGTCAATATCTGTGTTACCAACAGAGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 TGAAGTCAATATCTGTGTTACCAACAGAGATGATGATGATGATGATGATGATGATGAT 1140
OY 1141 CCATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 CCATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
OY 1201 TGTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 TGTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
OY 1261 TGTGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 TGTGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
OY 1321 GGGTGTAAAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1321 GGGTGTAAAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
OY 1381 ATAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1381 ATAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
OY 1441 ATATGGGCACTCCAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1441 ATATGGGCACTCCAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
OY 1501 GGGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1501 GGGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
OY 1561 TCTGTGAGAGAGGCTGTCTACTCAATTTGAGCTGCTCCACAGTTCCTTGTGTGT 1620
DB 1561 TCTGTGAGAGAGGCTGTCTACTCAATTTGAGCTGCTCCACAGTTCCTTGTGTGT 1620
OY 1621 CCATCAGAGCAGACAGAGGCTTGGCACTGATGATGATGATGATGATGATGATGATGAT 1680
DB 1621 CCATCAGAGCAGACAGAGGCTTGGCACTGATGATGATGATGATGATGATGATGATGAT 1680
OY 1681 GATACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
DB 1681 GATACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740

TITLE	JOURNAL	MEDLINE	PUBMED
transcript induced during dendritic cell differentiation	Immunogenetics 53 (12), 993-1001 (2002)	21901265	11904675
REFERENCE	Decker,J.W., Budhia,S., Angel,N.Z., Cooper,B.J., Clark,G.J., Hart,D.N.J. and Kato,M.	(bases 1 to 2677)	
AUTHORS	Hart,D.N.J. and Kato,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-Oct-2000) Dendritic Cell Research, Mater Medical Research Institute, Level 3, Aubigny Place, South Brisbane, Queensland 4101, Australia		
FEATURES	Location/Qualifiers		
SOURCE	1..2677 /organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/map="ip21"		
	/clone="211(1)B"		
	/cell_line="Hodgkin's disease cell line L428"		
	1..114 /note="derived from 5' RACE"		
CDS	369..1961 /note="adeno/homocysteine hydrolase-like protein; expressed in dendritic cells in blood, langerhans cells and dermal dendritic cells; mRNA increased during monocyte differentiation to monocyte-derived dendritic cells in the presence of GM-CSF and IL-4"		
	/codon_start=1		
	/product="S-adeno/homocysteine hydrolase-like protein"		
	/protein_id="AAI26869.1"		
	/db_xref="GI:16588687"		
	/translation="MSMPDAMPPLPGVEELKQAKIEDAEKYSEMATVTYKAPKOIOFADDDLEFTPEPTKTGRSLRSISQSDYSASAATDSDEVSREKQOTNSKSSSNFCVKNIKOAEFREREIEAODMSALILSRKAOGCEPLAGKIYGCHITROTAVLIERICLAGACRCSACNIYSTONEVAAILAEAVALPAMKESEDPMDCVMNMDGMANNLLDDGDLTTHRYKKTYNKRIKRGITYEBSYTGHRTXLSKAGLCVPMA NVNSTVRKFNLNYCCRESLTLDGLKTTDMFFEGKVVCIGEVSGCAALKALGAIAIVTEIDIPICALOMDGDFRRVKLVNEVDVTVCTGNKNVYVREHLDRKNSCIVCMASNTSIDVTSLRTELWERYRSQDVHWIMPDGRVLVLAAGRILNISTSVTPFLSTATTOTALILEYNAPGRYKDPLYLLPKMDEVASTLHPSFDAHLTETLTDOKRYGILKNKGEPKPYRYX"		
BASE COUNT	657 a 648 c 708 g 664 t		
ORIGIN			
Query Match	100.0%; Score 2563; DB 9; Length 2677;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 2563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
OY 1	GCGCGGGGCGAGCTGGAGCTCGGACTGTCTTGTTCTCTTGTCGCCACGTCGCT	60	
Dd 115	GCGCGGGGCGAGCTCGAGACTCGAGACTCTGCTTCTGTCTGTGTGCACCCTCGCT	174	
OY 61	GTCGGGTCGCTTGGGCGTCGGCGAACAGACAAGGGCTGGGGCCACACACTCGAAGCCGA	120	
Dd 175	GTCGGGTCGCTTGGGCGTCGGCGAACAGACAAGGGCTGGGGCCACACACTCGAAGCCGA	234	
OY 121	CGCAGCTCGACGACGAGGGCCGGCAGAGAAGGTTGGCGATCGCTGTCTGGAGGGCGCGCGC	180	
Dd 235	CGCAGCTCGACGACGAGGGCCGGCAGAGAAGGTTGGCGATCGCTGTCTGGAGGGCGCGCGC	294	
OY 181	GGGCGAGGCGGGGGCGGCCAGAGAGGGGAAAAGAGCGGGGGGGCGGGGTCAGCCGTCGGCC	240	
Dd 295	GGGCGAGGCGGGGGCGGCCAGAGAGGGGAAAAGAGCGGGGGGGCGGGGTCAGCCGTCGGCC	354	
OY 241	GGGCGGGGGGGGATGTGCGATGCGCTGACGCGCATGCCCTGCGCGGGGGCGGGAGAGAGC	300	
Dd 355	GGGCGGGGGGGGATGTGCGATGCGCTGACGCGCATGCCCTGCGCGGGGGCGGGAGAGAGC	414	
OY 301	TGAAGCAGAGCCAAAGAGATCGAGAGCGCGGAAATACTACTCTTCATGGCACCGTCACCA	360	
Dd 415	TGAAGCAGAGCCAAAGAGATCGAGAGCGCGGAAATACTACTCTTCATGGCACCGTCACCA	474	

QY	361	AGGGGCCCAAGAACCAATCCAGTTGGTGTATGACATCGGGAGTTCCACCAATTCOCCA	420
Db	475	AGGGGCCCAAGAACCAATCCAGTTGGTGTATGACATCGAGAGTTCCACCAATTCOCCA	534
QY	421	CCAAACTGGCCGAGATCTTTGTCTGCTGCATCTCAGTCCCTCCACTGCAGCTACA	480
Db	535	CCAAACTGGCCGAGATCTTTGTCTGCTGCATCTCAGTCCCTCCACTGCAGCTACA	594
QY	481	GTTTCAGTGCATCTTACACAGATAGCTCTGATGTAGTGAAGTTTCTCCCGAGAAAGCAGC	540
Db	595	GTTTCAGTGCATCTTACACAGATAGCTCTGATGTAGTGAAGTTTCTCCCGAGAAAGCAGC	654
QY	541	AAACCAACTCCAAAGGGGAGCAGCAATTTCTGTGGAAABAACATCAAGCAGAGAAATTTGG	600
Db	655	AAACCAACTCCAAAGGGGAGCAGCAATTTCTGTGGAAABAACATCAAGCAGAGAAATTTGG	714
QY	601	GACCGCGGAGATTGAGATTGACAGCAGAACATGTCTGTCTGTATTTCACTCAGGAAAC	660
Db	715	GACCGCGGAGATTGAGATTGACAGCAGAACATGTCTGTCTGTATTTCACTCAGGAAAC	774
QY	661	GTCCTCAGGGGGAGAAACCCCTTGGCTGTCTAAATAGTGGGCTGTACACATCAGAC	720
Db	775	GTCCTCAGGGGGAGAAACCCCTTGGCTGTCTAAATAGTGGGCTGTACACATCAGAC	834
QY	721	CCGAGACAGGGGATTTGATTGAGACACTCTGTGCCCTGAGGGGCTCAGAGCCGCTGGTCG	780
Db	835	CCGAGACAGGGGATTTGATTGAGACACTCTGTGCCCTGAGGGGCTCAGAGCCGCTGGTCG	894
QY	781	CTTGTAAACATCTACTCAACTCAGATGAAAGTAGCTGCAGCACTGGCTCAGCTGAGAGTTG	840
Db	895	CTTGTAAACATCTACTCAACTCAGATGAAAGTAGCTGCAGCACTGGCTCAGCTGAGAGTTG	954
QY	841	CATGTTCCGTTGGAAAGGGGAGTCAAGAAATGACTTGTGGTGTATTTGACCCGCTGTG	900
Db	955	CATGTTCCGTTGGAAAGGGGAGTCAAGAAATGACTTGTGGTGTATTTGACCCGCTGTG	1011
QY	901	TGAACATGATGGGTGGCAGGCCCAACATGATCCCGAATGAGGGGGAACCTTAACCCACT	960
Db	1015	TGAACATGATGGGTGGCAGGCCCAACATGATCCCGAATGAGGGGGAACCTTAACCCACT	1077
QY	961	GGGTTTATAAGAAATATCCAAACGTGTTTAAAGAAATCCAGGCATTTGTGAAAGAGAGCG	1020
Db	1075	GGGTTTATAAGAAATATCCAAACGTGTTTAAAGAAATCCAGGCATTTGTGAAAGAGAGCG	1133
QY	1021	TGAGCTGCTGTACACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGGCCA	1088
Db	1135	TGAGCTGCTGTGTACACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGGCCA	1199
QY	1141	CCATTTTGGATGGCCTGAAAGAGCACACAGATGTATGTTTGGTGGGAAACAAAGTGTGG	1200
Db	1255	CCATTTTGGATGGCCTGAAAGAGCACACAGATGTATGTTTGGTGGGAAACAAAGTGTGG	1311
QY	1201	TGTGTGGCTATGGTGTAGAGTAGGCAAGGGCTGTGTGCTCTCAAGCTCTTGGAGCAA	1266
Db	1315	TGTGTGGCTATGGTGTAGAGTAGGCAAGGGCTGTGTGCTCTCAAGCTCTTGGAGCAA	1377
QY	1261	TTGTGTACATTTACCGAAATGACCCCATCTGTGTGTGTGACAGGCTGTGATGGGTCCA	1322
Db	1375	TTGTGTACATTTACCGAAATGACCCCATCTGTGTGTGTGACAGGCTGTGATGGGTCCA	1433
QY	1321	GGGTGTAAAGCTAAATGAAAGTATCCGCGCAAGTCATGTCGTATATACTTGACAGGAA	1380
Db	1435	GGGTGTAAAGCTAAATGAAAGTATCCGCGCAAGTCATGTCGTATATACTTGACAGGAA	1499
QY	1381	ATPAAGATGTAGTACACGGGACACTTGGATCGCATGAAACAAAGTTGTATCGATGCA	1440
Db	1495	ATPAAGATGTAGTACACGGGACACTTGGATCGCATGAAACAAAGTTGTATCGATGCA	1555
QY	1441	ATATGGCCCACTCCAAACACAGAAATGATGTACACAGCCTTCGCACTTCGAGAGCTGACGT	1500

Db	1555	ATATGGGCACTCCACACAGAAATCGATGAGCACACCTCCGACCTCCGAGCTGAGCT	1614
Qy	1501	GGGAGGAGTACCTTTCAGAGTGACATCTGGCCGATGGGCAAGAGATGGCC	1566
Db	1615	GGGAGGAGTACCTTTCAGAGTGACATCTGGCCGATGGGCAAGAGATGGCC	1674
Qy	1561	TCTGGCAGAGGGTCTGTACTCAATTTGAGCTCTCCACAGTTCCACCTTGTCTGT	1620
Db	1675	TCTGGCAGAGGGTCTGTACTCAATTTGAGCTCTCCACAGTTCCACCTTGTCTGT	1733
Qy	1621	CCATCACAGCCACAACACAGGCTTGGCAGTGAATGAACTCTATAATGCAACCCGAGGGC	1688
Db	1735	CCATCACAGCCACAACACAGGCTTGGCAGTGAATGAACTCTATAATGCAACCCGAGGGC	1794
Qy	1681	GATACAGCAGATGTGTACTTCTCTTAAAGAAATGGATGATACGTTCCAGCTTGC	1740
Db	1795	GATACAGCAGATGTGTACTTCTCTTAAAGAAATGGATGATACGTTCCAGCTTGC	1855
Qy	1741	ATCTGCCATCATTTGATGTGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	1800
Db	1855	ATCTGCCATCATTTGATGTGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	1914
Qy	1801	GACTCAACAAAATGGGCGCATTAACCTATATTACAGTACTAATGACACTACTAC	1866
Db	1915	GACTCAACAAAATGGGCGCATTAACCTATATTACAGTACTAATGACACTACTAC	1974
Qy	1861	CAAGAGCAGTCCACCTGAAACACACACTCTAAAGAAATATTTTAAAGATACCTTAT	1920
Db	1975	CAAGAGCAGTCCACCTGAAACACACACTCTAAAGAAATATTTTAAAGATACCTTAT	2033
Qy	1921	TTTTCTTCTACTCCTTTCCTTGTGATTTTTTTCCTATAAATTCATCTGTGTTTTTCATC	1980
Db	2035	TTTTCTTCTACTCCTTTCCTTGTGATTTTTTTCCTATAAATTCATCTGTGTTTTTCATC	2094
Qy	1981	TGATTTCCAGTTCGAGAACACACAGGAATCTGCTATGGCTTTTGAATGAATA	2040
Db	2095	TGATTTCCAGTTCGAGAACACACAGGAATCTGCTATGGCTTTTGAATGAATA	2154
Qy	2041	GAAGTTCAGGGTCCCTCACTCTAGTCACTTAAAGAGATTTTACTCCCGCCAGCAGAA	2100
Db	2155	GAAGTTCAGGGTCCCTCACTCTAGTCACTTAAAGAGATTTTACTCCCGCCAGCAGAA	2214
Qy	2101	GGTGAATCTTCTTTTACATTTCTGGGCACTTACTTAAATTAAGTACTTATTAACA	2160
Db	2215	GGTGAATCTTCTTTTACATTTCTGGGCACTTACTTAAATTAAGTACTTATTAACA	2274
Qy	2161	GGAAATGCAAGGTACTCTCTGTGSGAANAATCTGAAATGTCTAATTCGCTTAAAGA	2220
Db	2275	GGAAATGCAAGGTACTCTCTGTGSGAANAATCTGAAATGTCTAATTCGCTTAAAGA	2334
Qy	2221	GCCCATTTTCTAGTCTCTAATAACAGTGCCTTTCATCTTTCAGAGAAGCAGGATGAT	2280
Db	2335	GCCCATTTTCTAGTCTCTAATAACAGTGCCTTTCATCTTTCAGAGAAGCAGGATGAT	2394
Qy	2281	ACCTACCCGGCAGGTAGTGAATGTGGGTGTCATGTTAAATTTCCCTTGAAGTTCA	2340
Db	2395	ACCTACCCGGCAGGTAGTGAATGTGGGTGTCATGTTAAATTTCCCTTGAAGTTCA	2454
Qy	2341	AGCCCTGTTTCTGGCTAAAGGGTATGTCCAGTTCAAGAGATGTGATATATGACATGG	2400
Db	2455	AGCCCTGTTTCTGGCTAAAGGGTATGTCCAGTTCAAGAGATGTGATATATGACATGG	2514
Qy	2401	CTTGTAAATCAGAGGCGCCACTTGATTTATATATACCCCTCCACACCTCCACCA	2460
Db	2515	CTTGTAAATCAGAGGCGCCACTTGATTTATATATACCCCTCCACACCTCCACCA	2574
Qy	2461	GACTTGCTCATTTTTCGAGTTTAACTAGTACACTCATTTTGAATTAATTTTCTC	2520
Db	2575	GACTTGCTCATTTTTCGAGTTTAACTAGTACACTCATTTTGAATTAATTTTCTC	2634
Qy	2521	TCTAGATTTATTTCTGTGTCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTT	2563

Db 2635 TCTAGATTTATTTCTGTGTCTCAAAAAAAAAAAAAAAAAAAAA 2677

RESULT 3
BC016942

LOCUS
DEFINITION
BC016942 2552 bp mRNA linear prt 09-NOV-2001
Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
MGC:21453 IMAGE:3450568, mRNA, complete cds.

ACCESSION
BC016942
VERSION
BC016942.1 GI:16877386

KEYWORDS
MGC.

SOURCE
Homo sapiens.
Homo sapiens.

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 2552)
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Strausberg, R.
Direct Submission
Submitted (05-NOV-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: ATCC
DNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mdickpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://lml.lnl.gov>
Series: IRAC Plate: 20 Row: a Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES

source

1. 2552
/organism="Homo sapiens"
/db_xref="LocusID:10768"
/db_xref="taxon:9606"
/clone="MGC:21453 IMAGE:3450568"
/tissue_type="Placenta, choriolealloma"
/clone_lib="NHL_MGC_10"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
338. 1840
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH16942.1"
/db_xref="GI:16877387"
/translation="MATVTKAPKKOIOFADDMQEFTEPKTKGRSLRSLSSISQSSYDS
YSSAASSTDSDDDEPREKQOTNSKNSCNFKIKOIEFPEREIEIODEMSALIS
LRRKROEKPLAGAKIVGCTHITAOAVLIEITCALGACRSASACIYISONVAAL
AAGVATPAFKMGSESDPMKIDRCVNMDCGMANNILDDGDLTHRYVKKPVEFKI
RGIVESVTEGVRHLKOLSKARKLCVPAKAVNDVTKOKDNITCCRESILDLKRTTD
VMEGKQVYVCGEYEVGKCCAAKALKAIAIYITIDIPICALQACMGRVYKLVNVI
ROYDVVITCTGNKNVYTRERHLRDKMNSCIVCMNGSMNEIDVTSILRTPETMERVRSO
VDHVIPTGKRVLLAEGRLLNLSCSTYPTFLSTATQALALILYNAPEGRYKOD
VYLPRKMDDEVASIHLPSPDAHLFELDDQAKYIGLNKNGFKPYVYR"

CDS

BASE COUNT 644 a 600 c 671 g 637 t

ORIGIN

Query Match 86.7%; Score 2222; DB 9; Length 2552;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2492; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 8 GCAGGTGCGAGCTGCGAGCTGCTGCTTCTGTCTTGTGCGCACCCTGCTGCTCCGCG 67

Db 1 GCAGGTGCGAGCTGCGAGCTGCTGCTTCTGTCTTGTGCGCACCCTGCTGCTCCGCG 60

QY 68 TGCCTTGGGCTGCGAGACAGACAGGCGTGGCCACAGCACTCAGAAAGCCAGCAGCT 127

Db 61 TGCCTTGGGCTGCGAGACAGACAGGCGTGGCCACAGCACTCAGAAAGCCAGCAGCT 120

QY 128 GCAGCGAGGCGCGCAGAGAGGCTGGGAGATCGGCTGTCGAGGCGCGCGCGGCGAGG 187

Db 121 CGAGCGAGGCGCGCAGAGAGGCTGGGAGATCGGCTGTCGAGGCGCGCGCGGCGAGG 180

QY 188 CGGCGGCGCGCAGAGAGGCGGAGAAAGAGCGGCGCGCGCGCGCGCTGCGCGCGCGCG 247

Db 181 CGGCGGCGCGCAGAGAGGCGGAGAAAGAGCGGCGCGCGCGCGCTGCGCGCGCGCG 240

QY 248 CGGCGGAGATGCTGATGCTGATGACGATGCGCGCTGCGCGCGCGCGCGCGAGCTGAAGA 307

Db 241 CGGCGGAGATGCTGATGCTGATGACGATGCGCGCTGCGCGCGCGCGCGAGCTGAAGA 300

QY 308 GCGCAAGAGATCGAGAGCGCGAGAGATGCTGATGCTGATGCGCGCGCGCGCGCGAGGCG 367

Db 301 GCGCAAGAGATCGAGAGCGCGAGAGATGCTGATGCTGATGCGCGCGCGCGCGAGGCG 360

QY 368 CAAGAGCAATTCAGATGCTGATGATGACGAGAGATGACCAATTCGCCACCAAAAC 427

Db 361 CAAGAGCAATTCAGATGCTGATGATGACGAGAGATGACCAATTCGCCACCAAAAC 420

QY 428 TGGCGGAAGATCTTGTCTGCTGATGATGACGATGACGATGCTGATGACGATGACGATGACG 487

Db 421 TGGCGGAAGATCTTGTCTGCTGATGATGACGATGACGATGACGATGACGATGACGATGACG 480

QY 488 TGCATCTTACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 547

Db 481 TGCATCTTACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540

QY 548 CTCGAGGAGGAGCAGCAGATTTCTGTGTGAAGAAATCAAGCAGGAGAAATTTGGAGCGG 607

Db 541 CTCGAGGAGGAGCAGCAGATTTCTGTGTGAAGAAATCAAGCAGGAGAAATTTGGAGCGG 600

QY 608 GAGAGTGAATGAGAGAGCAAGACATGCTGCTGATGATGATGATGATGATGATGATGATGATG 667

Db 601 GAGAGTGAATGAGAGAGCAAGACATGCTGCTGATGATGATGATGATGATGATGATGATGATG 660

QY 668 GGGGAGAGAGCCTTGGCTGTGTGTAAATTAATGAGGCTGTACACATACAGCCAGAC 727

Db 661 GGGGAGAGAGCCTTGGCTGTGTGTAAATTAATGAGGCTGTACACATACAGCCAGAC 720

QY 728 AGCGGTGTTGATGAGACACTGTGCGCTGGGGGCTCAGTGCGGCTGCTGTAA 787

Db 721 AGCGGTGTTGATGAGACACTGTGCGCTGGGGGCTCAGTGCGGCTGCTGTAA 780

QY 788 CATCTACTCACTCAGAAATGAATGATGCTGACACACTGCTGAGGCTGGAATTCAGATGT 847

Db 781 CATCTACTCACTCAGAAATGAATGATGCTGACACACTGCTGAGGCTGGAATTCAGATGT 840

QY 848 CGCTTGAAGGCGAGTCAAGAAATGACTTCTGCTGTATTTAGCCGCTGTGTGAACAT 907

Db 841 CGCTTGAAGGCGAGTCAAGAAATGACTTCTGCTGTATTTAGCCGCTGTGTGAACAT 900

QY 908 GATGCGGAGGAGCGCAGCAATGATCTGTGATGATGAGGAGAGCTTAACCACTGGGTTTA 967

Db 901 GATGCGGAGGAGCGCAGCAATGATCTGTGATGATGAGGAGAGCTTAACCACTGGGTTTA 960

QY 968 TAAGAAATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1027

Db 961 TAAGAAATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020

QY 1028 TGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTCTCCGCGCATGAAGCT 1087

Db 1021 TGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTCTCCGCGCATGAAGCT 1080

QY 1088 CAATGATCTGTATCAAAACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1147

D	1081	CAATATTCTGTTACCCAAACAAGATTGTATAACTGTACTGCTGCAGGAATTCATATTTT	1140
OY	1148	GGATGGCCTGAAGAGCACACAGANGTAGTGTTTTGGGAAACAAGTGGTGCTGTGG	1207
D	1141	GGATGGCCTGAAGAGCACACAGATGTGATGTTTGGTGGAAAACAAGTGGTGCTGTGG	1200
OY	1208	CTATGGTAGAGTAGGCAAGAAGGCGTGCCTGCTGCTGCCAAGCTCTTGGAGCAATGTCTTA	1267
D	1201	CTATGGTAGAGTAGGCAAGGCGTGCCTGCTGCTGCCAAGCTCTTGGAGCAATGTCTTA	1260
OY	1268	CATTACCGAATGACACCCTATCTGTGCTGTGCAGGCCCTGCATGGATGGTTCAAGGTGT	1327
D	1261	CATTACCGAATGACACCCTATCTGTGCTGTGCAGGCCCTGCATGGATGGTTCAAGGTGT	1320
OY	1328	AAAGCTAAATGAATGATCTATCCGGCAGTGCATGTCTGTAATATCTTGCACAGGAATTAAGA	1387
D	1321	AAAGCTAAATGAATGATCTATCCGGCAGTGCATGTCTGTAATATCTTGCACAGGAATTAAGA	1380
OY	1388	TGTACTGACAGGGGAGCACTTGGATGGCATGAAAAACAGTGTATGCTATGACATATGGG	1447
D	1381	TGTACTGACAGGGGAGCACTTGGATGGCATGAAAAACAGTGTATGCTATGACATATGGG	1440
OY	1448	CCACTCCAACACAGAAATGATGTGACACAGCCCTCCGCACTCCGGAGCTGACGTGGAGCG	1507
D	1441	CCACTCCAACACAGAAATGATGTGACACAGCCCTCCGCACTCCGGAGCTGACGTGGAGCG	1500
OY	1508	AGTACGTTCTCAGGTGGAGACCATGTCTCTGGCCAGATGGCAAAGAGTGTCTCTCTGGC	1567
D	1501	AGTACGTTCTCAGGTGGAGACCATGTCTCTGGCCAGATGGCAAAGAGTGTCTCTCTGGC	1560
OY	1568	AGAGGCTGCTACTACATCAATTTGAGCTGCCTCCACAGATGCCACCTTTGCTGTGCATCAC	1627
D	1561	AGAGGCTGCTACTACTCAATTTGAGCTGCCTCCACAGATGGCAAAGAGTGTCTCTCTGGC	1620
OY	1628	AGCCACAACACAGGCTTTGGCAGCTGATTAACCTATATATGACOC CGAGGGCGCATACAA	1687
D	1621	AGCCACAACACAGGCTTTGGCAGCTGATTAACCTATATATGACOC CGAGGGCGCATACAA	1680
OY	1688	GCAGAGTGTGTACTTGCTTCTCTTAAGAAAAATGATGAATACGTTGCCAGCTTGCATCTGCC	1747
D	1681	GCAGAGTGTGTACTTGCTTCTCTTAAGAAAAATGATGAATACGTTGCCAGCTTGCATCTGCC	1740
OY	1748	ATCATTTGATGGCCACCTTACAGAGCTGACAGATGACAAAGCAAAATATCGGACCTCAA	1807
D	1741	ATCATTTGATGGCCACCTTACAGAGCTGACAGATGACAAAGCAAAATATCGGACCTCAA	1800
OY	1808	CAAAATATGGGCATTCACAACTAATATTATACAGATCTATGACCATACTACCAAGAAC	1867
D	1801	CAAAATATGGGCATTCACAACTAATATTATACAGATCTATGACCATACTACCAAGAAC	1860
OY	1868	CAGTCACCTGACACACACACTCTAAGAATAATTTTTNAGATAACTTTATTTTCTTC	1927
D	1861	CAGTCACCTGACACACACACTCTAAGAATAATTTTTNAGATAACTTTATTTTCTTC	1920
OY	1928	TTACCTCTTTTCTCTTGATTTTTTTTCCCTAATATTCAATCTTGTTTTTCACTCATAT	1987
D	1921	TTACCTCTTTTCTCTTGATTTTTTTTCCCTAATATTCAATCTTGTTTTTCACTCATAT	1980
OY	1988	CCAAGTTTCGACAGACACACAGAGAACTTGCATCAGGCTTTAGATGAATAGAAATTC	2047
D	1981	CCAAGTTTCGACAGACACACAGAGAACTTGCATCAGGCTTTAGATGAATAGAAATTC	2040
OY	2048	AGGGGCTCTACCTAGTCACTAAGAGAGATTTTACTCCCCAGCCACAGAAAGGTAT	2107
D	2041	AGGGGCTCTACCTAGTCACTAAGAGAGATTTTACTCCCCAGCCACAGAAAGGTAT	2100
OY	2108	CTTCTCTTACCATTTCTGGGACCTTAAGCTTAATAGTACCTTAATTAACAGGAATG	2167
D	2101	CTTCTCTTACCATTTCTGGGACCTTAAGCTTAATAGTACCTTAATTAACAGGAATG	2159
OY	2168	CTAAGTACCTCTCTGGGAAACAATCTGCATGTCTAAATGCGCTTAAAGAGCCCAT	2227
D	2160	CTAAGTACCTCTCTGGGAAACAATCTGCATGTCTAAATGCGCTTAAAGAGCCCAT	2219

QY	2228	TCTTAGAGCTCAAAATCAAGTGCCTTTCACCTCTTACAGAAACAGAGGATGTAACCTAC	2287
Dp	2220	TCTTAGAGCTCAAAATCAAGTGCCTTTCACCTCTTACAGAAACAGAGGATGTAACCTAC	2279
QY	2288	CGGCGAGTATAGTTAGATGTGGGTGGTGCATGTAAATTTCCCTTGAAGTTCGAAGCCCTG	2347
Dp	2280	CGGCGAGTATAGTTAGATGTGGGTGGTGCATGTAAATTTCCCTTGAAGTTCGAAGCCCTG	2339
QY	2348	TTTCCGTGCGTAAAGGTGATGATGCCAGTTCAGAGATGTATATATGACATGGCTGTAA	24070
Dp	2340	TTTCCGTGCGTAAAGGTGATGATGCCAGTTCAGAGATGTATATATGACATGGCTGTAA	23959
QY	2408	AGATCAGAGAGGCCCACTGTGGATTTAATAGTATACCCCTCTCCACATCCACACAACTTGC	2467
Dp	2400	AGATCAGAGAGGCCCACTGTGGATTTAATAGTATACCCCTCTCCACATCCACACAACTTGC	24559
QY	2468	TCATTTTTCGATTTTAACTAGACTACACTATTT	2503
Dp	2460	TCATTTTTCGATTTTAACTAGACTACACTATTT	2495

RESULT 4	
BC007576	
LOCUS	2526 bp mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
ACCESSION	MGC:15558 IMAGE:3139729, mRNA, complete cds.
VERSION	BC007576
KEYWORDS	BC007576.1 GI:14043176
SOURCE	MGC.
ORGANISM	Homo sapiens.
REFERENCE	Homo sapiens.
AUTHORS	Eunayyola; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Mammalia; Primates; Carnivora; Hominiidae; Homo.
JOURNAL	1 (bases 1 to 2526)
	Strausberg,R.
	Direct Submission
	Submitted (10-MAY-2001) National Institutes of Health, Mammalian
	Gene Collection (MGC), Cancer Genomics Office, National Cancer
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
	USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: gcgabs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Field, Erin Garland, Ran Gulin, Iettitia Hsiao, Martin Krzywnski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Seediq, Jacqueline Schell, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorpe, Miraneda Tsai, Natasha van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Series: IRAL Plate: 22 Row: h Column: 21 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein. Location/Qualifiers 1. .2526 /organism="Homo sapiens" /db_xref="LocustID:10768" /db_xref="taxon:9606" /clone="MGC:15558 IMAGE:3139729" /tissue_type="Placenta, choriocarcinoma" /clone_id="NH.MGC_21"

```

/lab.host="DH10B-R"
/notes="Vector: pOMB7"
308..1810
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH07576.1"
/db_xref="GI:14043177"
/translation="MATVTKAPKKQIOPADMDQEFKFPFKTKGRSLSRSISQSTDS
YSSASVYDSDSEVSPREKQNTSKSSNFKVKNKQAEFGREIEIAEDMSALLS
LRRAGGERPLGAKIVGCTHTIATQAVLLETICALAQCRWSACNTYSTONENVAAL
.AEAGVAFAMKGESEDDFWWCIDRCVNMDOAMNLLDDGDLHMYVKKPKPYEKKI
RGIYESVTVHRLYOLSKAGKLCVPMNMDVSTYKQFNLXCCRSILDLKRTED
VNFGRKOVYVCYGEYFGKCCALKAIGALVYITEIDICALOCMRGFRVYKXNEVI
ROYDVITCTGNKNVYTRREHIDPMKNSCTYCNMHSHTBEIDYSLRPELWERYRSO
VDVWIPDKRYVLLLEGRLLNLSCTVPEFSTVATQALALIEYNAPEGRKOD
VYLPKMDIEVYASLHPSFADLITELTDQAKRYLGIKNKGFENYRY"
BASE COUNT      644 a      592 c      660 g      630 t
ORIGIN
Query Match      84.9%; Score 2177; DB 9; Length 2526;
Best Local Similarity 99.8%; Pident. No. 0;
Matches 2447; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

```

Db 736 CTGGTCTGCTTGTAACTACTACTCACTCAAGAAATGAAGTAGCTGACAGACTGGCTGAGGC 795
QY 833 TGGAGTTGCACTTTGCTTGGAAAGGCGAGTCAAGAAATGACTTTGGTGGTATTTGA 892
Db 796 TGGAGTTGCACTTTGCTTGGAAAGGCGAGTCAAGAAATGACTTTGGTGGTATTTGA 855
QY 893 CCGCTGTGCAATGAGTGGGTGGCGAGGCCAATGATCTGTGATGATGGGGGAGACT 952
Db 856 CCGCTGTGCAATGAGTGGGTGGCGAGGCCAATGATCTGTGATGATGGGGGAGACT 915
QY 953 AACCCACTGGGTTTATTAAGAAGTATCCAAACGTTTAAAGAATCCGAGGCAATTTGGA 1012
Db 916 AACCCACTGGGTTTATTAAGAAGTATCCAAACGTTTAAAGAATCCGAGGCAATTTGGA 975
QY 1013 AGAGAGCTGACTGGTGTTCACAGGCTGTATCACTCTCCAAAGCTGGAAAGCTTGTGT 1072
Db 976 AGAGAGCTGACTGGTGTTCACAGGCTGTATCACTCTCCAAAGCTGGAAAGCTTGTGT 1035
QY 1073 TCCGGCCATGAAGCTCAATGATTTCTGTACCAACAGAAAGTTGATTAAGTGTACTGCTG 1132
Db 1036 TCCGGCCATGAAGCTCAATGATTTCTGTACCAACAGAAAGTTGATTAAGTGTACTGCTG 1095
QY 1133 CCGAGAAATCCATTTTGGATGGCCCTGAAGAGGACCAACAGATGTGATGTTGGTGGAAACA 1192
Db 1096 CCGAGAAATCCATTTTGGATGGCCCTGAAGAGGACCAACAGATGTGATGTTGGTGGAAACA 1155
QY 1193 AGTGTGTGTGTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1252
Db 1156 AGTGTGTGTGTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1215
QY 1253 TGGAGCAATGTGTACATTTACCGAAATGACCCCATCTGTGCTGACAGGCTGATGGA 1312
Db 1216 TGGAGCAATGTGTACATTTACCGAAATGACCCCATCTGTGCTGACAGGCTGATGGA 1275
QY 1313 TGGGTTTACAGGTTGTTAAAGCTTAATGAAGTATCGGGAAGTGTGATGATGATGATGATGATGAT 1372
Db 1276 TGGGTTTACAGGTTGTTAAAGCTTAATGAAGTATCGGGAAGTGTGATGATGATGATGATGATGAT 1335
QY 1373 CACAGAAATTAAGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1432
Db 1336 CACAGAAATTAAGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1395
QY 1433 CGTATGCAATTAAGGAGGACCTCAACACAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1492
Db 1396 CGTATGCAATTAAGGAGGACCTCAACACAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1455
QY 1493 GCTGACGTGGGAGGAGTACCTTCTCAGGTGACCATGATGATGATGATGATGATGATGATGATGATGAT 1552
Db 1456 GCTGACGTGGGAGGAGTACCTTCTCAGGTGACCATGATGATGATGATGATGATGATGATGATGATGAT 1515
QY 1553 AGTTGCTCTCTGCGACAGAGGTCGTCTACTCAATTAAGTGTGCTCAGAGTCCACACTT 1612
Db 1516 AGTTGCTCTCTGCGACAGAGGTCGTCTACTCAATTAAGTGTGCTCAGAGTCCACACTT 1575
QY 1613 TGTTCCTCATCAACAGCCACACACAGAGGCTTTGGCACTGATAGAACTCTAATAAGCAC 1672
Db 1576 TGTTCCTCATCAACAGCCACACACAGAGGCTTTGGCACTGATAGAACTCTAATAAGCAC 1635
QY 1673 CGAGGGCGATTAAGAGAGATGTGTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATGATGAT 1732
Db 1636 CGAGGGCGATTAAGAGAGATGTGTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATGATGAT 1695
QY 1733 CAGTTGCTCATCTGATCATATTTGATGACCACTTACAGAGCTGACAGATGACCAAGCAAA 1792
Db 1696 CAGTTGCTCATCTGATCATATTTGATGACCACTTACAGAGCTGACAGATGACCAAGCAAA 1755
QY 1793 ATATCTGGAATCAACAAATTAAGGAGATTAAGGAGATTAAGGAGATTAAGGAGATTAAGGAGAT 1852
Db 1756 ATATCTGGAATCAACAAATTAAGGAGATTAAGGAGATTAAGGAGATTAAGGAGATTAAGGAGAT 1815
QY 1853 CATACTACCAAGAGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1912
Db 1816 CATACTACCAAGAGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1875

```


Db 503 GTTCACTGCTCCTACACAGATGCTCTGATGATGAGGTTTCCGCCGAGAGAAACGAC 562
 Oy 541 AACCAACTCCAAAGGCGAGCAGCAATTTCTGTGAAAGAAATCAAGAGCAGCAATTTG 600
 Db 563 AACCAACTCCAAAGGCGAGCAGCAATTTCTGTGAAAGAAATCAAGAGCAGCAATTTG 622
 Oy 601 GAGCGCGGAGATTTGAGATTTGACAGAGCAAGACATGTCTGCTGATTTCACTCAGAAAC 660
 Db 623 GAGCGCGGAGATTTGAGATTTGACAGAGCAAGACATGTCTGCTGATTTCACTCAGAAAC 682
 Oy 661 GTGCTCAGGAGGAGAGACCTTGGCTGTGTCTAAATAGTGGGCTGTACACATCAACAG 720
 Db 683 GTGCTCAGGAGGAGAGACCTTGGCTGTGTCTAAATAGTGGGCTGTACACATCAACAG 742
 Oy 721 CCCAGACACGCGTGTGATTTGAGACACTGTGTGCGCTGGGGGCTCAGTCCGCTGTCTG 780
 Db 743 CCCAGACACGCGTGTGATTTGAGACACTGTGTGCGCTGGGGGCTCAGTCCGCTGTCTG 802
 Oy 781 CTTGTAACTCTACTCACTCAGAAATGAGTGTGACAGCACTGGCTGAGGCTGAGTTG 840
 Db 803 CTTGTAACTCTACTCACTCAGAAATGAGTGTGACAGCACTGGCTGAGGCTGAGTTG 862
 Oy 841 CAGTGTTCCTTGGAAAGGCGAGTCAAGATGACTTCTGTGTGTATTTAGCCCTGTG 900
 Db 863 CAGTGTTCCTTGGAAAGGCGAGTCAAGATGACTTCTGTGTGTATTTAGCCCTGTG 922
 Oy 901 TGAACATGGATGGGCGAGGCGCAACATGATCTGTGATGGGGGAGACTTAACCCACT 960
 Db 923 TGAACATGGATGGGCGAGGCGCAACATGATCTGTGATGGGGGAGACTTAACCCACT 982
 Oy 961 GGGTTTATAGAAATATCCAAAGCTGTTTAAAGATCCGAGGCAATTTGGAAGAGAGG 1020
 Db 983 GGGTTTATAGAAATATCCAAAGCTGTTTAAAGATCCGAGGCAATTTGGAAGAGAGG 1042
 Oy 1021 TGAATGCTGTTACAGAGCTGTATGACTTCCAAAGCTGGGAAGCTGTGTCTCCGCCA 1080
 Db 1043 TGAATGCTGTTACAGAGCTGTATGACTTCCAAAGCTGGGAAGCTGTGTCTCCGCCA 1102
 Oy 1081 TGAACGATGATTTCTGTACCAACAGAAATTTGATTAATTTGATTAATTTGATTAATTT 1140
 Db 1103 TGAACGATGATTTCTGTACCAACAGAAATTTGATTAATTTGATTAATTTGATTAATTT 1162
 Oy 1141 CCAATTTGATGAGCTGTGAGAGCAACAGATGTATGTTGGTGGAAACAGTGGTG 1200
 Db 1163 CCAATTTGATGAGCTGTGAGAGCAACAGATGTATGTTGGTGGAAACAGTGGTG 1222
 Oy 1201 TGTGTGCTATGTTGAGTGTGAGGCAAGGCTGTGTGTCTCTCAAACTCTTGAAGCA 1260
 Db 1223 TGTGTGCTATGTTGAGTGTGAGGCAAGGCTGTGTGTCTCTCAAACTCTTGAAGCA 1282
 Oy 1261 TGTGTACATTTACGAAATGAGAACCCCATCTGTGCTGCAAGGCTGTGATGGGTTCA 1320
 Db 1283 TGTGTACATTTACGAAATGAGAACCCCATCTGTGCTGCAAGGCTGTGATGGGTTCA 1342
 Oy 1321 GGGTGTAAAGCTAAATGAATGATCCGCAAGTGTGATTAATTTGATTAATTTGATTAATTT 1380
 Db 1343 GGGTGTAAAGCTAAATGAATGATCCGCAAGTGTGATTAATTTGATTAATTTGATTAATTT 1402
 Oy 1381 ATAAGAAATGATGACACGAGGAGCACTTGTGATGATTAATTTGATTAATTTGATTAATTT 1440
 Db 1403 ATAAGAAATGATGACACGAGGAGCACTTGTGATGATTAATTTGATTAATTTGATTAATTT 1462
 Oy 1441 ATATGGGCGCTCCAAACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
 Db 1463 ATATGGGCGCTCCAAACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1522
 Oy 1501 GGGAGGAGTACGTTCTCAGGTGAGCAATGTCACTGTGCGAGATGGAACAGAGTTGTC 1560
 Db 1523 GGGAGGAGTACGTTCTCAGGTGAGCAATGTCACTGTGCGAGATGGAACAGAGTTGTC 1582
 Oy 1561 TCTGTGAGAGGCTGTCTACTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1620
 Db 1583 TCTGTGAGAGGCTGTCTACTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1642

Oy 1621 CCATCACAGCCACACACAGGCTTTGGCACTGATAGAACTCTATTAATGACCCGAGGAGC 1680
 Db 1643 CCATCACAGCCACACACAGGCTTTGGCACTGATAGAACTCTATTAATGACCCGAGGAGC 1702
 Oy 1681 GATACAGAGATGTGTACTTGTCTTCTTAAGAAATGATGATGATGATGATGATGATGATGATG 1740
 Db 1703 GATACAGAGATGTGTACTTGTCTTCTTAAGAAATGATGATGATGATGATGATGATGATGATG 1762
 Oy 1741 ATCTGCATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
 Db 1763 ATCTGCATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1822
 Oy 1801 GACTCAACAAAATGGGCACTTCAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860
 Db 1823 GACTCAACAAAATGGGCACTTCAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1882
 Oy 1861 CAAGGACAGTCCACCTGAAACACACACTCTAAGAAATATTTTAAAGATTAATTAATTAAT 1920
 Db 1883 CAAGGACAGTCCACCTGAAACACACACTCTAAGAAATATTTTAAAGATTAATTAATTAAT 1942
 Oy 1921 TTTCTTCTTCT 1980
 Db 1943 TTTCTTCTTCT 2002
 Oy 1981 TCATTAATCCAAAGTCTGACAGACACACAGCAAGTCTGCTCAATGCTCTTAAGATAATA 2040
 Db 2003 TCATTAATCCAAAGTCTGACAGACACACAGCAAGTCTGCTCAATGCTCTTAAGATAATA 2062
 Oy 2041 GAAGTTGAGGCT 2100
 Db 2063 GAAGTTGAGGCT 2122
 Oy 2101 GGTGATCT 2160
 Db 2123 GGTGATCT 2181
 Oy 2161 GGAATATGCTAAGTACTCTCTCTGGAACAATCTGCAATCTGAATCTGAATCTGAATCTGAAT 2220
 Db 2182 GGAATATGCTAAGTACTCTCTCTGGAACAATCTGCAATCTGAATCTGAATCTGAATCTGAAT 2241
 Oy 2221 GGCATTTCTTCT 2280
 Db 2242 GGCATTTCTTCT 2301
 Oy 2281 ACCTACCCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
 Db 2302 ACCTACCCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2361
 Oy 2341 AGCCCTGTTCT 2400
 Db 2362 AGCCCTGTTCT 2421
 Oy 2401 CTTGTTAAGATCAGAGAGCCCATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2460
 Db 2422 CTTGTTAAGATCAGAGAGCCCATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2481
 Oy 2461 GACTGTGCTATTTTGAAGTTTAACTAGACTACACTTAAT 2503
 Db 2482 GACTGTGCTATTTTGAAGTTTAACTAGACTACACTTAAT 2524

RESULT 6
 HSU82761
 LOCUS
 DEFINITION
 Homo sapiens S-adenosyl homocysteine hydrolase homolog (Kpivone)
 mRNA, complete cds.
 ACCESSION
 U82761
 VERSION
 U82761.1 GI:2852124
 KEYWORDS
 Homo sapiens.
 Homo sapiens
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 2258)
AUTHORS Cleaver,J.E., Afzal,V., Feeney,L., McDowell,M., Sadinski,W., Volpe,J.P.G., Busch,D.B., Coleman,D.M., Ziffer,D.W., Yu,Y., Nagasawa,H. and Little,J.B.
TITLE Increased ultraviolet sensitivity and chromosomal instability related to p53 function in the xeroderma pigmentosum variant
JOURNAL Cancer Res. 59 (5), 1102-1108 (1999)
MEDLINE 99168517
PUBMED 10070969
REFERENCE 2 (bases 1 to 2258)
AUTHORS Volpe,J.P.G., McDowell,M., Jostes,R.F., Afzal,V., Sadinski,W., Traub,B.J., Legerski,R. and Cleaver,J.E.
TITLE Complementation of chromosomal instability in the xeroderma pigmentosum variant by a gene on human chromosome 1 with homology to S-adenosyl homocysteine hydrolase
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2258)
AUTHORS Volpe,J.P.G., McDowell,M. and Cleaver,J.E.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1996) Dermatology, UCSF, 3rd and Parnassus, Box 0750, San Francisco, CA 94143, USA
FEATURES
source 1. 2258
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/cell_type="fibroblasts"
/tissue_type="skin"
1. 2258
/gene="XPVKona"
48. 1550
/note="Xeroderma pigmentosum variant"
/product="S-adenosyl homocysteine hydrolase homolog"
/protein_id="AAC01960.1"
/translation="MATVTKAPKKQIQIADMOEFTFPPTGRRSRISISQSTDS
YSSASTSDSDDEVSPREKQOTNSKSNVCNINIKOAEGRREIETAEODMSALIS
LRRKAOEKLPAKIVGCTHITQAVLITLICALGACMSACNIEIESTOREVAAL
AEGAVFAFKGSESDFWNCIDRCVNMNDQAMHLLDDGDLTFHWVKKPKPNVKKI
RGIVBESVTGVHHLXOLSKAGKLCVPMANNVDSITKOFDNLCCRESIDLGKRTTD
VMFGKOVVCGTIGEVGKCCALALGALVITLIDPICALQACMDGFRVKKLEVI
ROVDVITCTGNKNVTVREHLDRKNSCIYCNMHSNTEIDVTSLRPELMEWRVSO
VDHVIWPDCKRVVLAEGRLINLSCSTVPEVLSITATTOALALIELNNAEGRYKOD
VLLPKMDRYVSLHLPSEDAHLTELTDQAKVLGINKNPFKNVRY"

BASE COUNT 603 a 513 c 542 g 600 t
ORIGIN

Query Match 72.8%; Score 1866; DB 9; Length 2258;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2086; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 298 AGCTGAAGCAGGCCAAGAGATCGAGACGCCGAGAAATCTCTTCATGGCCACCGCTA 357
DB 1 AGCTGAAGCAGGCCAAGAGATCGAGACGCCGAGAAATCTCTTCATGGCCACCGCTA 60
QY 358 CCAAGGCGCCCAAGAGCAATCCAGTTTGTGATGACATGAGAGATTACCAATTC 417
DB 61 CCAAGGCGCCCAAGAGCAATCCAGTTTGTGATGACATGAGAGATTACCAATTC 120
QY 418 CCAAACTGGCCGGAAGATCTTGTCTCGTCGATCTACAGTCCCTCACTGACAGCT 477
DB 121 CCAAACTGGCCGGAAGATCTTGTCTCGTCGATCTACAGTCCCTCACTGACAGCT 180
QY 478 ACAATTGAGTCTCTACACAGATAGCTGTATGATGAGGTTTCTCCCGAGAGAAC 537
DB 181 ACAATTGAGTCTCTACACAGATAGCTGTATGATGAGGTTTCTCCCGAGAGAAC 240
QY 538 AGCAACCACTCCAAAGGAGCAGCAGCAATTTCTGTGAAGAAACATCAAGCAGAGAT 597
DB 241 AGCAACCACTCCAAAGGAGCAGCAGCAATTTCTGTGAAGAAACATCAAGCAGAGAT 300

QY 598 TTGAGACCGCGGAGATTGAGATTGACAGACAGACATGCTGCTGTGATTCATCAGAGA 657
DB 301 TTGAGACCGCGGAGATTGAGATTGACAGACAGACATGCTGCTGTGATTCATCAGAGA 360
QY 658 AAGTGTCAAGGGGGAAGAGCCCTTGGTGTGCTAAATAGTGGGCTGTACACATCA 717
DB 361 AAGTGTCAAGGGGGAAGAGCCCTTGGTGTGCTAAATAGTGGGCTGTACACATCA 420
QY 718 CAGCCCAAGACGCGGTGTGATTGATGAGACACTGTGTCCTGGGCGCTAGTCCCGTGT 777
DB 421 CAGCCCAAGACGCGGTGTGATTGATGAGACACTGTGTCCTGGGCGCTAGTCCCGTGT 480
QY 778 CTGCTTTAATCTACTCACTCAGATGAATGATAGTGTGACACACTGTGCTGAGCTGAG 837
DB 481 CTGCTTTAATCTACTCACTCAGATGAATGATAGTGTGACACACTGTGCTGAGCTGAG 540
QY 838 TTGAGTGTGCGGTGGAAGGGCGAGTGAAGAGATCTTGCTGGTGGTATGACCGCT 897
DB 541 TTGAGTGTGCGGTGGAAGGGCGAGTGAAGAGATCTTGCTGGTGGTATGACCGCT 600
QY 898 GTGTGAACATGATGAGGTGGCAGGCCCAACATGATCTGATGATGGGAGACTTAACC 957
DB 601 GTGTGAACATGATGAGGTGGCAGGCCCAACATGATCTGATGATGGGAGACTTAACC 660
QY 958 ACTGGGTTTATAGACATGATCCAAACGTTTAAAGATCCGAGCATTTGTGAAGAGA 1017
DB 661 ACTGGGTTTATAGACATGATCCAAACGTTTAAAGATCCGAGCATTTGTGAAGAGA 720
QY 1018 GCGTGAATGCTGTCACAGGCTGATGCTGCTCCAAAGCTGGGAAGCTGCTGTCGG 1077
DB 721 GCGTGAATGCTGTCACAGGCTGATGCTGCTCCAAAGCTGGGAAGCTGCTGTCGG 780
QY 1078 CCATGAACGTCATGATCTGTTTACCAAGAGAGAGTTGATTAATCTGCTGCTGCGAG 1137
DB 781 CCATGAACGTCATGATCTGTTTACCAAGAGAGAGTTGATTAATCTGCTGCTGCGAG 840
QY 1138 AATCCATTTTGGATGCGCTGAAGAGAGACACAGATGTGATTTGGTGGAAACAAGTGG 1197
DB 841 AATCCATTTTGGATGCGCTGAAGAGAGACACAGATGTGATTTGGTGGAAACAAGTGG 900
QY 1198 TGGGTGTGGTATGATGAGGTAGGCAAGGCGCTGCTGCTGCTGCTGCTGCTGAG 1257
DB 901 TGGGTGTGGTATGATGAGGTAGGCAAGGCGCTGCTGCTGCTGCTGCTGCTGAG 960
QY 1258 CAATTGCTCATTAACGGAATTCAGCCATCTGCTCTGAGGCGCTGATGATGGGT 1317
DB 961 CAATTGCTCATTAACGGAATTCAGCCATCTGCTCTGAGGCGCTGATGATGGGT 1020
QY 1318 TCAGGGTGTAAAGCTAAATGATCATCGGCAAGTGTGATTAATCTGTCACAG 1377
DB 1021 TCAGGGTGTAAAGCTAAATGATCATCGGCAAGTGTGATTAATCTGTCACAG 1080
QY 1378 GAATATGATGTGTATGACACAGGACACTTGGATCGGATGAAAAACAGTTTATCGTAT 1437
DB 1081 GAATATGATGTGTATGACACAGGACACTTGGATCGGATGAAAAACAGTTTATCGTAT 1140
QY 1438 GCAATATGGGCACTCCAAACAGAAATCGATGTGACAGGCTCGGCACTCGGAGCTGA 1497
DB 1141 GCAATATGGGCACTCCAAACAGAAATCGATGTGACAGGCTCGGCACTCGGAGCTGA 1200
QY 1498 CGTGGAGCAGATGCTTCTCAGGTGACATGTCATCTGGCCAGATGGCAACAGATTG 1557
DB 1201 CGTGGAGCAGATGCTTCTCAGGTGACATGTCATCTGGCCAGATGGCAACAGATTG 1260
QY 1558 TCTCTCTGGCAGAGGGTCTCTACTCAATTTGAGTGTCTCCACAGTTCCACCTTTGTT 1617
DB 1261 TCTCTCTGGCAGAGGGTCTCTACTCAATTTGAGTGTCTCCACAGTTCCACCTTTGTT 1320
QY 1618 TGTTCATCAGACGACCAACACAGGCTTTGGACGTGATGAATCTATATATGACCCGAGG 1677
DB 1321 TGTTCATCAGACGACCAACACAGGCTTTGGACGTGATGAATCTATATATGACCCGAGG 1380

QY 854 GAAGGCGAGTCAGAGATGACTTCTGTGTGATGACCCCTGTGTGAATGATG 913
 DB 803 GAAAGGCGAGTCAGAGATGACTTCTGTGTGATGACCCCTGTGTGAATGATG 862
 QY 914 GTGCGAGCCCAACATGATCTGTGATGATGGGAGACCTTAACCCATGGTTATAGAA 973
 DB 863 GTGCGAGCCCAACATGATCTGTGATGATGGGAGACCTTAACCCATGGTTATAGAA 922
 QY 974 GTATCCAAACGTGTTTAAAGATGCCAGGATGTGGAAGAGAGCGTACTGTGTTCA 1033
 DB 923 GTATCCAAACGTGTTTAAAGATGCCAGGATGTGGAAGAGAGCGTACTGTGTTCA 982
 QY 1034 CAGGCTGTATAGCTCTTCCAAAGCTGGGAGCTGTGTCTCCGGCATTAAGTCAATCA 1093
 DB 983 CAGGCTGTATAGCTCTTCCAAAGCTGGGAGCTGTGTCTCCGGCATTAAGTCAATCA 1042
 QY 1094 TTTCTGTATCCAAAGAGATTTGATTAATGATCTGTGCGAGATTCATTTTGGATGG 1153
 DB 1043 TTTCTGTATCCAAAGAGATTTGATTAATGATCTGTGCGAGATTCATTTTGGATGG 1102
 QY 1154 CCTGGAAGAGACCAAGATGTGATGTTTGTGGGAACAGTGGTGGTGTGGCTATGG 1213
 DB 1103 CCTGGAAGAGACCAAGATGTGATGTTTGTGGGAACAGTGGTGGTGTGGCTATGG 1162
 QY 1214 TGAGGTAGGCAAGGGCTGTGTCTGTCTCAAAAGCTCTTGGAGCAATGTCTACTTAC 1273
 DB 1163 TGAGGTAGGCAAGGGCTGTGTCTGTCTCAAAAGCTCTTGGAGCAATGTCTACTTAC 1222
 QY 1274 CGAATGAGACCCCATCTGTGTCTGTCAGAGCCCTGCATGATGGGTTCAGGGTGAAGCT 1333
 DB 1223 CGAATGAGACCCCATCTGTGTCTGTCAGAGCCCTGCATGATGGGTTCAGGGTGAAGCT 1282
 QY 1334 AAATGAAGTCATCCGGCAAGTCGATGTGTATTAATCTTGCACAGGAATTAAGATTAAGT 1393
 DB 1283 AAATGAAGTCATCCGGCAAGTCGATGTGTATTAATCTTGCACAGGAATTAAGATTAAGT 1342
 QY 1394 GACACGGAGAGACTTGGATCCGATTAAGAAAACAGTTGATCGATTAATGAGGCGACATC 1453
 DB 1343 GACACGGAGAGACTTGGATCCGATTAAGAAAACAGTTGATCGATTAATGAGGCGACATC 1402
 QY 1454 CAACACAGAAATGATGATGACAGAGCTCCGACCTCCGAGCTGAGCTGGAGGAGATAGG 1513
 DB 1403 CAACACAGAAATGATGATGACAGAGCTCCGACCTCCGAGCTGAGCTGGAGGAGATAGG 1462
 QY 1514 TTTCTAGGTGAGACATGATCTGTGCGAGATGCGCAAGAGTGTCTCTGCGACAGAGG 1573
 DB 1463 TTTCTAGGTGAGACATGATCTGTGCGAGATGCGCAAGAGTGTCTCTGCGACAGAGG 1522
 QY 1574 TCTGTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTTCTGTCTCATACAGCCAC 1633
 DB 1523 TCTGTCTACTCAATTTGAGCTGCTCCACAGTTCCACCTTTGTTCTGTCTCATACAGCCAC 1582
 QY 1634 AACACAGGCTTTGGACATGATAGAACTCTATTAATGACCCCGAGGGGCGATTACAGCAGA 1693
 DB 1583 AACACAGGCTTTGGACATGATAGAACTCTATTAATGACCCCGAGGGGCGATTACAGCAGA 1642
 QY 1694 TGTGTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1753
 DB 1643 TGTGTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1702
 QY 1754 TGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAACAAAAA 1813
 DB 1703 TGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAACAAAAA 1762
 QY 1814 TGAGGCAATTAACCTTAATTTTACAGATTAATGAGCAATGACCAAGCAAGCAAGTCC 1873
 DB 1763 TGAGGCAATTAACCTTAATTTTACAGATTAATGAGCAATGACCAAGCAAGCAAGTCC 1822
 QY 1874 ACCGGAACACACACTCTAAAGAAATATTTTAAAGATTAATTTTAACTTCTCTACTC 1933
 DB 1823 ACCGGAACACACACTCTAAAGAAATATTTTAAAGATTAATTTTAACTTCTCTACTC 1882
 QY 1934 CTTTCCCTCTGATTTTCTTCTAATTAATTTCTCTGTTTCTTCAATTCATTAATCAAGT 1993

DB 1883 CTTTCCCTCTGATTTTCTTCTAATTAATTTCTGTTTCTTCAATTCATTAATCAAGT 1942
 QY 1994 TCTGACAGACCAACAGAACTTGTCTTACAGCTCTTTTGAATGAATAGTTCAGGGTC 2053
 DB 1943 TCTGACAGACCAACAGAACTTGTCTTACAGCTCTTTTGAATGAATAGTTCAGGGTC 2002
 QY 2054 CTTCTACTCTAGTCACTAAAGAGATTTTACTCCCGACCCAGAAAGGTATTTCTCTC 2113
 DB 2003 CTTCTACTCTAGTCACTAAAGAGATTTTACTCTCTCCAGCCAGAAAGGTATTTCTT -TC 2061
 QY 2114 TTTACCAATTTCTGGGACTTATAGCTTATTAAGTACTTATTAAGAGAAATGCTAAG 2173
 DB 2062 TTTACCAATTTCTGGGACTTATAGCTTATTAAGTACTTATTAAGAGAAATGCTAAG 2121
 QY 2174 TACCTCTCTGTGGAACAAATCTGCAATGCTAAATCGCCTTAAAGAGCCATTTCTTAG 2233
 DB 2122 TACCTCTCTGTGGAACAAATCTGCAATGCTAAATCGCCTTAAAGAGCCATTTCTTAG 2181
 QY 2234 CTGCTGAATCAGTGTCTTCTTCACTTCTTCAAGAGAGAGGATGATACCCGGCAG 2293
 DB 2182 CTGCTGAATCAGTGTCTTCTTCACTTCTTCAAGAGAGAGGATGATACCCGGCAG 2241
 QY 2294 GTAGGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2353
 DB 2242 GTAGGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2301
 QY 2354 GCGTAAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2413
 DB 2302 GCGTAAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2361
 QY 2414 GAGAGCCCACTTGAATTAATGATAGTATGAGCCCTCTTCACTCCACCAAGATGCTCATTT 2473
 DB 2362 GAGAGCCCACTTGAATTAATGATAGTATGAGCCCTCTTCACTCCACCAAGATGCTCATTT 2421
 QY 2474 TTTGAGTTTAACTAATCACTCACTTAT 2503
 DB 2422 TTTGAGTTTAACTAATCACTCACTTAT 2451

RESULT 8
 AF035319
 LOCUS 1369 bp mRNA linear PRI 07-JUL-2000
 DEFINITION Homo sapiens clone 23931 mRNA, partial cds.
 ACCESSION AF035319
 VERSION AF035319.1 GI:2661082
 KEYWORDS FLI.CDNA.
 SOURCE
 ORGANISM Homo sapiens.
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Anderson, B., Wentland, M. A., Ricalfente, J. Y., Liu, W. and Gibbs, R. A.
 TITLE A 'double adaptor' method for improved shotgun library construction
 JOURNAL Anal. Biochem. 236 (1), 107-113 (1996)
 MEDLINE 96207227
 PUBMED 8619474
 REFERENCE
 AUTHORS Yu, W., Anderson, B., Morley, R. C., Muzny, D. M., Ding, Y., Liu, W., Ricalfente, J. Y., Wentland, M. A., Lennom, G. and Gibbs, R. A.
 TITLE Large-scale concatenation cDNA sequencing
 JOURNAL Genome Res. 7 (4), 353-358 (1997)
 MEDLINE 9110174
 PUBMED 9110174
 REFERENCE
 AUTHORS Yu, W., Sarginson, J. and Gibbs, R. A.
 TITLE Direct Submission
 JOURNAL Submitted (20-NOV-1997) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA
 FEATURES
 source 1..1369
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

```

/note="similar to S-adenosylhomocysteine hydrolase"
/codon_start=1
/protein_id="AAB88189.1"
/db_xref="GI:3661083"
/translation="PRAMGYSVEVGKCCALALGAIYITEDIPICALADAGDFG
TWAKLAEVLRQVDVITTCGKNNVYIEHLDNRNSCTVCSMGHSNITDIYS
IVERRRQADHWIWDGKRWVLLAEERLLNLSCTVFCVLSIATTAQALILY
PEGRYKQDVYLLPKMDEYVASLHPSFAHLTELTDQAKYLGILNNNGPKPYIR

```

Query Match	42.1%	Score 1080;	DB 9;	Length 1369;
Best Local Similarity	99.8%	Pred. No. 0;		
Matches 1300;	Conservative	0;	Mismatches 2;	Indels 1;
				Gaps 1.

QY	1201	TGTGTGCTATGTAGGTAGGACGAAGGGCTGCTGTGCTCTCTCAAAAGCTTTGGAGCAA	1260
Db	11	TGTGTGCTATGTAGGTAGGACGAAGGGCTGCTGTGCTCTCTCAAAAGCTTTGGAGCAA	70
QY	1261	TTGTCTACATTAACCGAAATCCAGCCCACTGTGTGCTCTGAGAGCCGCGCATGGAATGGGTTCA	1320
Db	71	TTGTCTACATTAACCGAAATCCAGCCCACTGTGTGCTCTGAGAGCCGCGCATGGAATGGGTTCA	130
QY	1321	GGGTGGTAAAGCTAATGTAAGTCAATCCGGCAAGTGGATGCTAATTAATTCACACAGAA	1380
Db	131	GGGTGGTAAAGCTAATGTAAGTCAATCCGGCAAGTGGATGCTAATTAATTCACACAGAA	190
QY	1381	ATAGATGATGATGACACGGGAGCACTTGGATGCGATGAAACAGTTGATCGTATGCA	1440
Db	191	ATAGATGATGATGACACGGGAGCACTTGGATGCGATGAAACAGTTGATCGTATGCA	250
QY	1441	ATATGGGCACTCCCAACACAGAAATCGAATGATACACAGCTCCGCACTCCGGAGCTGACGT	1500
Db	251	ATATGGGCACTCCCAACACAGAAATCGAATGATACACAGCTCCGCACTCCGGAGCTGACGT	310
QY	1501	GGGAGCGAGTACGTTCTCAGGTGAGACCAATGTCATCTGGCCAGATGGCAACAAGTTGTCC	1560
Db	311	GGGAGCGAGTACGTTCTCAGGTGAGACCAATGTCATCTGGCCAGATGGCAACAAGTTGTCC	370
QY	1561	TCCTGGCAGAGGGTGTCTACTCAATTTGAGTGTCCACAGTTCCACCTTTGTTCTGT	1620
Db	371	TCCTGGCAGAGGGTGTCTACTCAATTTGAGTGTCCACAGTTCCACCTTTGTTCTGT	430
QY	1621	CCATACACAGCCACACACAGAGCTTTGGCAGCTATAGAACTCTATATATGACCCGAGGGCC	1680
Db	431	CCATACACAGCCACACACAGAGCTTTGGCAGCTATAGAACTCTATATATGACCCGAGGGCC	490
QY	1681	GATACACAGAGATGTTACTTGTCTCCCTAAAGAAATGGATGAATAAACGTTGGCAGGTTGC	1740
Db	491	GATACACAGAGATGTTACTTGTCTCCCTAAAGAAATGGATGAATAAACGTTGGCAGGTTGC	550
QY	1741	ATTCG ⁴ CCATCATTTTATGATCCCACTTATAGAGCTACAGATGACCAAGCAAAATATCTGG	1800
Db	551	ATTCGCCATCATTTTATGATCCCACTTATAGAGCTACAGATGACCAAGCAAAATATCTGG	610
QY	1801	GACTCACAAAAATGGCCATTCAAACTAATTTATTACAGATCTATAGGACCATCTAC	1860
Db	611	GACTCACAAAAATGGCCATTCAAACTAATTTATTACAGATCTATAGGACCATCTAC	670
QY	1861	CAGAGCACAGTCCACTGAACACACACACTCTAAAGAAATATTTTAAATAAATCTTAT	1920
Db	671	CAGAGCACAGTCCACTGAACACACACACTCTAAAGAAATATTTTAAATAAATCTTAT	730

QY	1921	TTTCTTCTACTCCCTTCCTCTGTTGTTTTTTCCTAATATTCATCTCTGTTTTTCAATC	1980
Db	731	TTTTCTCTTACTCCCTTCCTTCCTCTTGATTTTTTTCCTATATTTTCAATCTCTGTTTTTTCATC	790
QY	1981	TCATATTCACAAGTCTCTGCAGACACACAGCAACTGCTTCATGGCTCTTTAGATGAATA	2040
Db	791	TCATATTCACAAGTCTCTGCAGACACACAGCAACTGCTTCATGGCTCTTTAGATGAATA	850
QY	2041	GAAGTTCAAGGGTCCCTCACTACTAGTCACTAAGAAAGATTTTACTCCCCCAGCCAGAA	2100
Db	851	GAAGTTCAAGGGTCTCTCACTACTAGTCACTAAGAAAGATTTTACTCTCCAGCCAGAA	910
QY	2101	GGTGGTTCTTCCTTTACTACATTTCTGGGGACCTTAGTATTAATAGACCTTTATTA	2160
Db	911	GGTGGTTCTTT-TCCTTTACCATTTCTGGGGACCTTAGTATTAATAGACCTTTATTA	969
QY	2161	GGAATTCGTAAAGTACCTTCTCTGTGGAAACATCTGCATGTCTAAATCGCCTTAAAA	2220
Db	970	GGAATTCGTAAAGTACCTTCTCTGTGGAAACATCTGCATGTCTAAATCGCCTTAAAA	1029
QY	2221	GCCCATTTCTTAGCTGCTGGAATACAGTCTCTTTCACTTTCTTCAGAGAAGAGGATG	2280
Db	1030	GCCCATTTCTTAGCTGCTGGAATACAGTCTCTTTCACTTTCTTCAGAGAAGAGGATG	1089
QY	2281	ACCTACCCGCGAGAGGTATGATGTGGGTGGTGCATGTATTAATTTCCCTTAGAATTCCA	2340
Db	1090	ACCTACCCGCGAGAGGTATGATGTGGGTGGTGCATGTATTAATTTCCCTTAGAATTCCA	1149
QY	2341	AGCCCTGTTTCCCTCGGTAAAGGTGTATGTCACGTTCCAGAGATGTATTAATGAGCATGG	2400
Db	1150	AGCCCTGTTTCCCTCGGTAAAGGTGTATGTCACGTTCCAGAGATGTATTAATGAGCATGG	1209
QY	2401	CTTGTTAGATCAGAGAGGCCCACTTGGATTTATAGTATAGCCCTCCCTCCACTCCACCA	2460
Db	1210	CTTGTTAGATCAGAGAGGCCCACTTGGATTTATAGTATAGCCCTCCCTCCACTCCACCA	1269
QY	2461	GACTGTCATTTTTCGAGTTTATTAACATAGACTACTCTATTT	2503
Db	1270	GACTGTCATTTTTCGAGTTTATTAACATAGACTACTCTATTT	1312

RESULT	9
LOCUS	BC003631
DEFINITION	Homo sapiens, similar to S-adenosylhomocysteine hydrolase-like 1,
ACCESSION	BC003631
VERSION	BC003631.1 GI:13177702

KEYWORDS	SOURCE	ORGANISM	REFERENCE
		Homo sapiens.	
		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
			1 (bases 1 to 1323)

AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (26-FEB-2001) National Institutes of Health, Mammalian

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk

Email: c9apbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
contact: amadansystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kelleman and Anuradha Madan

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 3 Row: D Column: 6.

FEATURES
 source location/Qualifiers

1. 1323
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3010755"
 /tissue_type="Muscle, rhabdomyosarcoma"
 /clone_id="NH_MGC_17"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 <1..611
 /codon_start=3
 /product="Similar to S-adenosylhomocysteine hydrolase-like
 1"

CDS

BASE COUNT 365 a 303 c 271 g 384 t
 ORIGIN
 Query Match 40.48; Score 1035; DB 9; Length 1323;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1246 AAGCTCTGGAGCAATTCACATTAACGAAATGACCCCAATCGTCGACAGCCG 1305
 10 AAGCTCTGGAGCAATTCACATTAACGAAATGACCCCAATCGTCGACAGCCG 69
 1306 GCATGATGGGTGAGGTGGTAAAGTAAATGATGATCCGCGCAAGTCGATGTG 1365
 70 GCATGATGGGTGAGGTGGTAAAGTAAATGATGATCCGCGCAAGTCGATGTG 129
 1366 TAACCTGCACAGGAATAGAAATGATGATGACACGGGAGCAGCTTGATGAT 1425
 130 TAACTTGCACAGGAATAGAAATGATGATGACACGGGAGCAGCTTGATGAT 189
 1426 GTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1485
 190 GTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 249
 1486 CTCGGAGCTGACGTGGAGGAGGAGTACGTTTCAGGTGGACCATGTCATGTC 1545
 250 CTCGGAGCTGACGTGGAGGAGGAGTACGTTTCAGGTGGACCATGTCATGTC 309
 1546 GCACAGAGTGTCTCTCTGCGAGAGGTCGTCTACTCAATTTGAGCTGTCAC 1605
 310 GCACAGAGTGTCTCTCTGCGAGAGGTCGTCTACTCAATTTGAGCTGTCAC 369
 1606 CCACCTTGTCTCTCTCTGCGAGAGGTCGTCTACTCAATTTGAGCTGTCAC 1665
 370 CCACCTTGTCTCTCTCTGCGAGAGGTCGTCTACTCAATTTGAGCTGTCAC 429
 1666 ATGACCCGAGGGGAGGATACAGAGAGTGTACTTCTCTCTCAAGAAATGAT 1725
 430 ATGACCCGAGGGGAGGATACAGAGAGTGTACTTCTCTCTCAAGAAATGAT 489
 1726 ACGTTGCCAGCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1785
 490 ACGTTGCCAGCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 549
 1786 AAGCAAAATATCTGGGACCAACAAATGAGGCAATTAACCTAATTAATTA 1845
 550 AAGCAAAATATCTGGGACCAACAAATGAGGCAATTAACCTAATTAATTA 609
 1846 AATGACCATATCTACCAAGGACGACGTCACCTGACACACACACTTAAGAA 1905
 610 AATGACCATATCTACCAAGGACGACGTCACCTGACACACACACTTAAGAA 669
 1906 TAAGATTAATCTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1965

DB 670 TAAATTAATCTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 729
 QY 1966 TCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2025
 DB 730 TCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 789
 QY 2026 TCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2085
 DB 790 TCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 849
 QY 2086 CCCCAGCCAGGAGGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2145
 DB 850 CTCACAGCCAGGAGGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 908
 QY 2146 GGTACCTTATTAACAGGAATGCTAAGTACCTTCTTCTTCTTCTTCTTCT 2205
 DB 909 GGTACCTTATTAACAGGAATGCTAAGTACCTTCTTCTTCTTCTTCTTCTTCT 968
 QY 2206 AATGACCTTAAAGAGCCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2265
 DB 969 AATGACCTTAAAGAGCCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1028
 QY 2266 AGAAGCAGGAGTGTACCTTACCCGAGGATGATGATGATGATGATGATGAT 2325
 DB 1029 AGAAGCAGGAGTGTACCTTACCCGAGGATGATGATGATGATGATGATGAT 1088
 QY 2326 CCCTTAGAGTTCACAGCCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2385
 DB 1089 CCCTTAGAGTTCACAGCCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1148
 QY 2386 GTATTAATGAGCATGCTTGTAAATGATGATGATGATGATGATGATGAT 2445
 DB 1149 GTATTAATGAGCATGCTTGTAAATGATGATGATGATGATGATGATGAT 1208
 QY 2446 CCTCACCTCCACAGGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2503
 DB 1209 CCTCACCTCCACAGGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1266

RESULT 10
 AR068569
 LOCUS AR068569 2226 bp DNA linear PAT 29-SEP-1999
 DEFINITION Sequence 2 from patent US 5854023.
 ACCESSION AR068569
 VERSION AR068569.1 GI:6000776
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1. (bases 1 to 2226)
 AUTHORS Hallman, D.L., Corley, N.C., Lai, P. and Shah, P.
 TITLE Polynucleotides encoding human S-adenosyl-5-homocysteine hydrolase
 derived from bladder
 JOURNAL Patent: US 5854023-A 2 29-DEC-1998;
 FEATURES
 source location/Qualifiers
 1..2226
 /organism="unknown"

BASE COUNT 649 a 455 c 435 g 675 t 12 others
 ORIGIN
 Query Match 21.6%; Score 553; DB 6; Length 2226;
 Best Local Similarity 99.6%; Pred. No. 7.7e-311;
 Matches 773; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1728 GTTGACAGCTGACATGTCATTTGATGACCACTTACAGAGTGCAGATGACCA 1787
 DB 1 GTTGACAGCTGACATGTCATTTGATGACCACTTACAGAGTGCAGATGACCA 60
 QY 1788 GCACAAATATCTGGGACCAACAAATGAGGCAATTAACCTAATTAATTA 1847
 DB 61 GCACAAATATCTGGGACCAACAAATGAGGCAATTAACCTAATTAATTAATTA 120

Matches	773:	Conservative	0:	Mismatches	2:	Indels	1:	Gaps	1:
Best Local Similarity 99.6%; Pred. No. 7,76-311;									
QY	1728	GTTCGACGCTTGCACTCGCATCAATTTGATGCGCACCTTACAGAGCTGACAGATGACCAA	1787						
Db	1	GTTCGACGCTTGCACTCGCATCAATTTGATGCGCACCTTACAGAGCTGACAGATGACCAA	60						
QY	1788	GCAAAATATCTGGGACTCAACAAAATGSGCATTCACAAACCTAATTTATTACAGATACAA	1847						
Db	61	GCAAAATATCTGGGACTCAACAAAATGSGCATTCACAAACCTAATTTATTACAGATACAA	120						
QY	1848	TGACCATATCTACACAGGACGATCCACTGAAACACACACTCTAAGAAGATATTTTAA	1907						
Db	121	TGACCATATCTACACAGGACGATCCACTGAAACACACACTCTAAGAAGATATTTTAA	180						
QY	1908	AGATACTTTTATTTTCTTCTTACTGCTTCTCTCTGATTTTTCATATTTTCATTC	1967						
Db	181	AGATACTTTTATTTTCTTCTTACTGCTTCTCTCTGATTTTTCATATTTTCATTC	240						
QY	1968	TGTGTTTTTTCATCTCATTTATCCAAATTTCTGACAGACACACAGAGACTGCTTCAGGCTC	2027						
Db	241	TGTGTTTTTTCATCTCATTTATCCAAATTTCTGACAGACACACAGAGACTGCTTCAGGCTC	300						
QY	2028	TTTAAATGAAATAGAAAGTTCAAGGTCCCTCACTCATGACATCAAAAGAGATTTTACTCC	2087						
Db	301	TTTAAATGAAATAGAAAGTTCAAGGTCCCTCACTCATGACATCAAAAGAGATTTTACTCT	360						
QY	2088	CCGACGCCAGAAAGGTGATTTCTTCTTTTACCATTTCTGGGGACTTATGCTTAAATAGG	2147						
Db	361	CCGACGCCAGAAAGGTGATTTCTTCTTTTACCATTTCTGGGGACTTATGCTTAAATAGG	419						
QY	2148	TACCTTATTTAAACAGAAATGCTAAGGTACTTCTCTGAGAACAACTTGCATATGCTTAA	2207						
Db	420	TACCTTATTTAAACAGAAATGCTAAGGTACTTCTCTGAGAACAACTTGCATATGCTTAA	479						
QY	2208	TGCGCTTAAAGAGGCCATTTCTTATGCTGCTGAATACAGTGTCTTTCACCTCTTCAGAG	2267						
Db	480	TGCGCTTAAAGAGGCCATTTCTTATGCTGCTGAATACAGTGTCTTTCACCTCTTCAGAG	539						
QY	2268	AAGCAGGAGTGTGTACTTACCACCGGACAGATGTTAGATGTGGTGGTGCATGTTAAATTC	2327						
Db	540	AAGCAGGAGTGTGTACTTACCACCGGACAGATGTTAGATGTGGTGGTGCATGTTAAATTC	599						
QY	2328	CTTAGAAGTCTCAAGCCCTGTTTCCCTGCGTAAAGTGTATCTCCAGTTACAGATGTGT	2387						
Db	600	CTTAGAAGTCTCAAGCCCTGTTTCCCTGCGTAAAGTGTATCTCCAGTTACAGATGTGT	659						
QY	2388	ATAAAGAGATGCGTGTGTAAGATACAGAGGCCCACTTGATTTATAGATAGCCCTTCC	2447						
Db	660	ATAAAGAGATGCGTGTGTAAGATACAGAGGCCCACTTGATTTATAGATAGCCCTTCC	719						
QY	2448	TCCACTTCCACACAGACTTGCTCATTTTTCGATTTTAACTAGACTACACTATT	2503						
Db	720	TCCACTTCCACACAGACTTGCTCATTTTTCGATTTTAACTAGACTACACTATT	775						
RESULT 12									
AL772411 81378 bp DNA linear PRI 16-AUG-2002									
LOCUS Human DNA sequence from clone Rpl1-180N18 on chromosome 1, complete									
DEFINITION sequence.									
ACCESSION AL772411 AC025865 AL360171									
VERSION AL772411.1 GI:21534613									
KEYWORDS HFG.									
SOURCE human.									
ORGANISM Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.									
REFERENCE 1 (bases 1 to 81378)									
AUTHORS Martin, S.									
TITLE Direct Submission									
JOURNAL Submitted (16-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,									
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:									

AX366273/c AX366273 292 bp DNA linear PAT 15-FEB-2002
LOCUS Sequence 40 from Patent WO0206317.
DEFINITION AX366273
ACCESSION AX366273
VERSION AX366273.1 GI:18697699
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Mitcham, J.L., King, G.E., Algate, P.A., Fling, S.P., Retter, M.W.,
Fanger, G.R., Reed, S.G., Vedvick, T.S., Carter, D., Hill, P. and
Albone, E.
TITLE Compositions and methods for the therapy and diagnosis of ovarian
cancer
JOURNAL Patent: WO 0206317-A 40 24-JAN-2002;
CORIAX CORPORATION (US)
FEATURES Location/Qualifiers
source 1..292
BASE COUNT 98 a 68 c 60 g 66 t
ORIGIN
Query Match 9.28; Score 237; DB 6; Length 292;
Best Local Similarity 100.0%; Pred. No. 4e-126;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2267 GAAGCAGGATGTACCTACCCGCGAGGTAGATGTGGTGGTGCATGTATTTTC 2326
DB 276 GAAGCAGGAGGTGTACCTACCCGCGAGGTAGATGTGGTGGTGCATGTATTTTC 217
QY 2327 CCTAGAAGTTCCAAAGCCCTGTTCCCTGCTAAAGGTGTATGTCAGTTCAAGATGTG 2386
DB 216 CCTAGAAGTTCCAAAGCCCTGTTCCCTGCTAAAGGTGTATGTCAGTTCAAGATGTG 157
QY 2387 TATPATAGCATGTGCTTTAGATCAGAGGCCACCTGATTTATAGTATAGCCCTTC 2446
DB 156 TATPATAGCATGTGCTTTAGATCAGAGGCCACCTGATTTATAGTATAGCCCTTC 97
QY 2447 CTCACCTCCACACAGACTGCTCATTTTTCAGATTTTAACTAGACTACACTCTATT 2503
DB 96 CTCACCTCCACACAGACTGCTCATTTTTCAGATTTTAACTAGACTACACTCTATT 40

Search completed: April 20, 2003, 21:15:35
Job time : 6890 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 17:23:45 ; Search time 543 Seconds
(without alignments)
10629.598 Million cell updates/sec

Title: US-09-782-051-1

Perfect score: 2563

Sequence: 1 ggcgcggcgagctcgagct.....aaaaaaaaaaaaaaaaaaaaa 2563

Scoring table: OLIGO_MNC

Gapop 60.0 ; Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2563	100.0	2563	19	AAV28617
2	2178	85.0	3634	23	ABV23195
3	2178	85.0	3634	23	ABV29032
4	1522	59.4	3616	20	AAV73924
5	570	22.2	721	21	AAV16004
6	553	21.6	2226	20	AAV73925
7	422	16.5	504	23	ABV44141
8	370	14.4	636	21	AA80766
9	330	12.9	376	24	ABN94533

10	320	12.5	349	21	AAA44806
11	237	9.2	292	21	AA69730
12	237	9.2	292	24	ABW72624
13	224	8.7	553	24	ABL79045
14	190	7.4	1877	18	AAV06354
15	168	6.6	384	24	ABN93848
16	118	4.6	312	23	ABV35453
17	118	4.6	435	23	ABV14215
18	117	4.6	326	21	AA28464
19	95	3.7	278	24	ABW73754
20	82	3.2	328	23	ABV05046
21	69	2.7	433	21	AA21046
22	66	2.6	138	16	AA25056
23	60	2.3	60	24	ABN33274
24	26	1.0	207	22	AAH71190
25	26	1.0	259	18	AA791303
26	26	1.0	259	18	AAV00423
27	26	1.0	315	23	ABV35453
28	26	1.0	352	23	ABV14369
29	26	1.0	451	22	AA109877
30	26	1.0	451	22	AA118549
31	26	1.0	452	22	AA118548
32	26	1.0	492	22	AA109876
33	26	1.0	863	18	AA791302
34	26	1.0	987	24	ABN98493
35	26	1.0	1165	24	AA23915
36	26	1.0	1789	24	ABQ88146
37	25	1.0	90	23	ABV35515
38	25	1.0	187	24	ABK39436
39	25	1.0	232	23	AAV19107
40	25	1.0	241	23	ABV08717
41	25	1.0	284	24	ABL87534
42	25	1.0	353	21	AA677649
43	25	1.0	373	22	AA68727
44	25	1.0	396	22	AA697700
45	25	1.0	410	22	AA191275

ALIGNMENTS

RESULT 1	AAV28617	standard; cDNA; 2563 BP.
XX	AAV28617;	
AC	21-AUG-1998	(first entry)
XX		
DT		
DE		
XX		
XX		
KW	DD4b9.3 enzyme; S-adenosyl-L-homocysteine hydrolase-type activity;	
KW	AHC-type activity; stimulation; inhibition; dendritic cell; cancer;	
KW	autoimmune disease; transplantation; ss.	
XX		
OS	Homo sapiens.	
XX		
FT		
FT		
FT		
FT		
PN	W09814562-A1.	
XX		
PD	09-APR-1998.	
XX		
PF	06-OCT-1997;	97MO-NZ00133.
XX		
PR	04-OCT-1996;	96NZ-0299507.
XX		
PA	(HART/) HART D-N J.	
XX		
PI	Hart DNT;	


```

Db 1801 GACTCAACAAAAAGGCGCATCTACACCTAATTTATACAGTACTAAAGACCACTAC 1860
Qy 1861 CAAGACACAGTCCACCTGAACACACACTCTAAGAAATATTTTAAATTAATTTAT 1920
Db 1861 CAAGACACAGTCCACCTGAACACACACTCTAAGAAATATTTTAAATTAATTTAT 1920
Qy 1921 TTTCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980
Db 1921 TTTCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980
Qy 1981 TCAATATCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2040
Db 1981 TCAATATCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2040
Qy 2041 GAAGTTCAGGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Db 2041 GAAGTTCAGGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Qy 2101 GGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
Db 2101 GGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
Qy 2161 GGAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2220
Db 2161 GGAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2220
Qy 2221 GCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280
Db 2221 GCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280
Qy 2281 ACCTACCCGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
Db 2281 ACCTACCCGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
Qy 2341 AGCCCTGTTTCCCTGCTGAAGAGTGTGATGATGATGATGATGATGATGATGATGAT 2400
Db 2341 AGCCCTGTTTCCCTGCTGAAGAGTGTGATGATGATGATGATGATGATGATGATGAT 2400
Qy 2401 CTGTGTAGATCAGAGGCGCCACTTGATTTATAGTATAGCCCTTCCCTCCACTCCCA 2460
Db 2401 CTGTGTAGATCAGAGGCGCCACTTGATTTATAGTATAGCCCTTCCCTCCACTCCCA 2460
Qy 2461 GACTGCTCATTTTTCAGATTTTAACTAGACTACACTATTTTAACTTTTGTCC 2520
Db 2461 GACTGCTCATTTTTCAGATTTTAACTAGACTACACTATTTTAACTTTTGTCC 2520
Qy 2521 TCTAGATTTTCTCTGTCGCAAAAAA 2563
Db 2521 TCTAGATTTTCTCTGTCGCAAAAAA 2563

```

RESULT 2

ABV23195
ID ABV23195 standard; cDNA: 3634 BP.

```

AC ABV23195;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 23186.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX

```

```

PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
PI WPI; 2001-662795/76.
XX
DR
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer; useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 4168-4169; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or incidence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;
XX
Query Match 85.0%; Score 2178; DB 23; Length 3634;
Best Local Similarity 99.8%; Pred. No. 0; Mismatches 4; Indels 1; Gaps 1;
Matches 2498; Conservative 0;
XX
Qy 1 GGCAGGCGGAGAGTGGAGCTGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db 71 GGCAGGCGGAGAGTGGAGCTGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 130
Qy 61 GTCCGCTGCTTGGGCTGCGGACAGACAGAGCGTGGCCACAGCAGCTCAGAACCGCA 120
Db 131 GTCCGCTGCTTGGGCTGCGGACAGACAGAGCGTGGCCACAGCAGCTCAGAACCGCA 190
Qy 121 CGCAGCTGAGAGCGGCGCGGAGAGAGTGGGAGTGGCGCTGCGGAGGCGCGGCG 180
Db 191 CGCAGCTGAGAGCGGCGCGGAGAGAGTGGGAGTGGCGCTGCGGAGGCGCGGCG 250
Qy 181 GGGCAGGCGGCGGCGCGGAGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 240
Db 251 GGGCAGGCGGCGGCGCGGAGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 310
Qy 241 GGGCGGCGGCGGAGTGTGATGCTGACGCGATGCGCGCTGCGCGGCGGCGGAGAGC 300
Db 311 GGGCGGCGGCGGAGTGTGATGCTGACGCGATGCGCGCTGCGCGGCGGCGGAGAGC 370
Qy 301 TGAAGCAGGCGCAAGAGATCGAGAGCGCGGAGAGTACTCTTCAATGGCCAGCTACCA 360
Db 371 TGAAGCAGGCGCAAGAGATCGAGAGCGCGGAGAGTACTCTTCAATGGCCAGCTACCA 430
Qy 361 AGGCGCCCAAGAGCAATCCAGTTGCGATGACATGAGAGTACACCAATTCGCCCA 420
Db 431 AGGCGCCCAAGAGCAATCCAGTTGCGATGACATGAGAGTACACCAATTCGCCCA 490
Qy 421 CCAAACTGCGCGAAGATCTTGTCTGCTGATCTCAGATCTCAGAGTCTCCAGTACAGCTACA 480
Db 491 CCAAACTGCGCGAAGATCTTGTCTGCTGATCTCAGATCTCAGAGTCTCCAGTACAGCTACA 550

```

```

481 GTTCAGCTGCACTCTACAGATAGCTCTGATGATGAGTTTCTCCCGAGAGACAGC 540
551 GTTCAGCTGATCTCTACACATAGCTCTGATGATGAGTTTCTCCCGAGAGACAGC 610
541 AACCACTCCAAAGGAGGAGCGAGCAATTTCTGTGAGAACATCAAGAGAGAGATTGG 600
611 AACCACTCCAAAGGAGGAGCGAGCAATTTCTGTGAGAACATCAAGAGAGAGATTGG 670
601 GAGCGCGGAGATTGAGATTGCGAGCAAGACATGCTGCTGTGATTTCACTCAGAGAAC 660
671 GAGCGCGGAGATTGAGATTGCGAGCAAGACATGCTGCTGTGATTTCACTCAGAGAAC 730
661 GTGCTCAGGGGAGAGAGCCCTGCTGCTGCTAAATAGTGGGCTGTACACATCAG 720
731 GTGCTCAGGGGAGAGAGCCCTGCTGCTGCTAAATAGTGGGCTGTACACATCAG 790
721 CCCAGAGAGGGGTGTGATTGAGACACTGTGGCCCTGGGGCTAGAGCCGCTGGTGG 780
791 CCCAGAGAGGGGTGTGATTGAGACACTGTGGCCCTGGGGCTAGAGCCGCTGGTGG 850
781 CTGTAACTCTACTCACTCACTGAGATGAGATGAGTGCAGCACTGGCTGAGGCTGGAGTGG 840
851 CTGTAACTCTACTCACTCACTGAGATGAGATGAGTGCAGCACTGGCTGAGGCTGGAGTGG 910
841 CAGGTTCGCTTGGAGAGGCGAGTCAGAGATGACTTCTGTGTGATTTGACCGCTGTG 900
911 CAGGTTCGCTTGGAGAGGCGAGTCAGAGATGACTTCTGTGTGATTTGACCGCTGTG 970
901 TGAACATGATGGGTGGAGGCGCAACATGATCCTGATGATGGGGGAGACTTAACCCACT 960
971 TGAACATGATGGGTGGAGGCGCAACATGATCCTGATGATGGGGGAGACTTAACCCACT 1030
961 GGGTTTAAAGAACTATCCAAACGTTTAAAGAACTCCGAGGCAATTTGGAAGAGAGCG 1020
1031 GGGTTTAAAGAACTATCCAAACGTTTAAAGAACTCCGAGGCAATTTGGAAGAGAGCG 1090
1021 TGACTGCTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGGCCA 1080
1091 TGACTGCTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGGCCA 1150
1081 TGAACATGATGGGTGGAGGCGCAACATGATCCTGATGATGGGGGAGACTTAACCCACT 1140
1151 TGAACATGATGGGTGGAGGCGCAACATGATCCTGATGATGGGGGAGACTTAACCCACT 1210
1141 CCATTTTGGATGGGCTGTAAGAGAGACACAGATGATGTTGGTGGGAACAAAGTGTGG 1200
1211 CCATTTTGGATGGGCTGTAAGAGAGAGACACAGATGATGTTGGTGGGAACAAAGTGTGG 1270
1201 TGTTGGCTATGTTGAGGTAGGCAAGGCTGCTGTGCTCTCAAAAGCTCTTGGAGCAA 1260
1271 TGTTGGCTATGTTGAGGTAGGCAAGGCTGCTGTGCTCTCAAAAGCTCTTGGAGCAA 1330
1261 TTGCTCATATTACCGGAATGACCCCATCTGTGCTCTGACAGGCTGCATGATGGGTTCA 1320
1331 TTGCTCATATTACCGGAATGACCCCATCTGTGCTCTGACAGGCTGCATGATGGGTTCA 1390
1321 GGGTGTAAAGCTAAATGAAATCCGCGCAAGTGCATGCTTAATAAATTGGACAGGAA 1380
1391 GGGTGTAAAGCTAAATGAAATCCGCGCAAGTGCATGCTTAATAAATTGGACAGGAA 1450
1381 ATAGATGATGATGACAGCGGAGACATGATCGCATGAAAAAGTGTATGATGCA 1440
1451 ATAGATGATGATGACAGCGGAGACATGATCGCATGAAAAAGTGTATGATGCA 1510
1441 ATATGGGCACTCCACACAGAAATGATGATGACAGGCTCCGCACTCCGAGCTGAGCT 1500
1511 ATATGGGCACTCCACACAGAAATGATGATGACAGGCTCCGCACTCCGAGCTGAGCT 1570
1501 GGGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
1571 GGGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1630
1561 TCCCTGAGAGAGGCTGTCTACTCAATTTGAGCTCTCCACAGTCTCCACCTTTGTTCTGT 1620

1631 TCCCTGAGAGAGGCTGTCTACTCAATTTGAGCTCTCCACAGTCTCCACCTTTGTTCTGT 1690
1621 CCATTCACAGGCGACACACAGAGCTTTGGCACTGATGATGATGATGATGATGATGATGAT 1680
1691 CCATTCACAGGCGACACACAGAGCTTTGGCACTGATGATGATGATGATGATGATGATGAT 1750
1681 GATTCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1751 GATTCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1810
1741 ATCTGCATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
1811 ATCTGCATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1870
1801 GACTCAACAAATAATGGGCACTTCAACCTAATTTATACAGATGATGATGATGATGATGATGAT 1860
1871 GACTCAACAAATAATGGGCACTTCAACCTAATTTATACAGATGATGATGATGATGATGATGAT 1930
1861 CAAGAGCAGTCCACACTGAAACACACACTCTAAGAAATATTTTAAAGATTAATCTTTAT 1920
1931 CAAGAGCAGTCCACACTGAAACACACACTCTAAGAAATATTTTAAAGATTAATCTTTAT 1990
1921 TTCTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980
1991 TTCTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2050
1981 TCATTTATCCAAAGTCTGACAGACACACAGAACTGCTCAATGGCTCTTATAGTAATA 2040
2051 TCATTTATCCAAAGTCTGACAGACACACAGAACTGCTCAATGGCTCTTATAGTAATA 2110
2041 GAAATTCAGAGGCTCCCTCACTCTACTACTACTAAGAAAGATTTTACTCCCGCAGCCAGAA 2100
2111 GAAATTCAGAGGCTCCCTCACTCTACTACTACTAAGAAAGATTTTACTCCCGCAGCCAGAA 2170
2101 GGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
2171 GGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2229
2161 GGAATATCTAAGGATACCTCTCTCTGGAACAACTGCAATGCTAATGCTCTTAAAGA 2220
2230 GGAATATCTAAGGATACCTCTCTCTGGAACAACTGCAATGCTAATGCTCTTAAAGA 2289
2221 GCCCATTTCTTAGCTGTGAATGATGCTCTTCACTCTCTCAGAGAGAGGATGCT 2280
2290 GCCCATTTCTTAGCTGTGAATGATGCTCTTCACTCTCTCAGAGAGAGGATGCT 2349
2281 ACCTACCCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
2350 ACCTACCCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2409
2341 AGCCCTGTTCTGCGTAAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
2410 AGCCCTGTTCTGCGTAAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2469
2401 CTGTGTAAGATCAGAGAGGCCCACTTGATTAATAGTATAGCCCTTCTCCACTCCACCA 2460
2470 CTGTGTAAGATCAGAGAGGCCCACTTGATTAATAGTATAGCCCTTCTCCACTCCACCA 2529
2461 GACTTGTCTATTTTTCAGATTTTAACTAGACTACACTCTATT 2503
2530 GACTTGTCTATTTTTCAGATTTTAACTAGACTACACTCTATT 2572

RESULT 3
ABV29032
ID ABV29032 standard; cdna; 3634 bp.
XX
AC ABV29032;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cdna 29023.
```



```

Db 1451 ATAGAGATGATGACACGGGAGACCTGGATCCATGAAAAACAGTTGATCGATGCA 1510
1441 ATATGGGCCACTCCACACAGAAATGATGAGACAGCCTCCGACCTCGGAGCTGACGT 1500
1511 ATATGGGCCACTCCACACAGAAATGATGAGACAGCCTCCGACCTCGGAGCTGACGT 1570
1501 GGGAGGAGATGACCTCTCAGGTGAGCAGTGTCTGAGCCAGATGGCAAGAGTTGTCC 1560
1571 GGGAGGAGATGACCTCTCAGGTGAGCAGTGTCTGAGCCAGATGGCAAGAGTTGTCC 1630
1561 TCCTGGCAGAGGGTCCGTCTACTCAATTTGAGCTGCTCCACAGTTCCACTTTGTCTGT 1620
1631 TCCTGGCAGAGGGTCCGTCTACTCAATTTGAGCTGCTCCACAGTTCCACTTTGTCTGT 1690
1621 CCATCAGACGACACACAGAGCTTTGGCAGCTAGATGAACTATATATGACCCGAGGGGC 1680
1691 CCATCAGACGACACACAGAGCTTTGGCAGCTAGATGAACTATATATGACCCGAGGGGC 1750
1681 GATACAAAGCAGATGCTGCTGCTCTCTTAAGAAATGATGATGATGCTGCTGCTGCTGC 1740
1751 GATACAAAGCAGATGCTGCTGCTCTCTTAAGAAATGATGATGATGCTGCTGCTGCTGC 1810
1741 ATCTGCATATTTGATGCTGCTGCTCTCTTAAGAAATGATGATGATGCTGCTGCTGCTGC 1800
1811 ATCTGCATATTTGATGCTGCTGCTCTCTTAAGAAATGATGATGATGCTGCTGCTGCTGC 1870
1801 GACTCAACAAAAATGGGCACTTCAAACTTATTTATTCAGATGATGATGATGATGATGATGAT 1860
1871 GACTCAACAAAAATGGGCACTTCAAACTTATTTATTCAGATGATGATGATGATGATGATGAT 1930
1861 CAAGAGCAGCTCCACCTGACACACACACACTTAAAGAAATTTTAAAGATGATGATGAT 1920
1931 CAAGAGCAGCTCCACCTGACACACACACACTTAAAGAAATTTTAAAGATGATGATGAT 1990
1921 TTTCTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980
1991 TTTCTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2050
1981 TCATATCAAGATGCTGACAGACACACAGAACTGCTTCACTGCTTCTTATGATGATGAT 2040
2051 TCATATCAAGATGCTGACAGACACACAGAACTGCTTCACTGCTTCTTATGATGATGAT 2110
2041 GAAGTGAAGGCTCCCTCAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
2111 GAAGTGAAGGCTCCCTCAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2170
2101 GGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
2171 GGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2229
2161 GGAATATCTAAGTACCTCTCTCTGGAACAATGCAATGCTTAAATGCTTAAAGA 2220
2230 GGAATATCTAAGTACCTCTCTCTGGAACAATGCAATGCTTAAATGCTTAAAGA 2289
2221 GCCCATTTCTAAGTCTGTAATGATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280
2290 GCCCATTTCTAAGTCTGTAATGATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2349
2281 ACCTACCCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
2350 ACCTACCCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2409
2341 AGCCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400
2410 AGCCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2469
2401 CTGTATAGATCAGAGGAGCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
2470 CTGTATAGATCAGAGGAGCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2529
2461 GACTGCTCATTTTGGAGTTTAACTAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2503

```

```

Db 2530 GACTGCTCATTTTTCGAGTTTAACTAGACTACTCTATT 2572
RESULT 4
AAV73924
ID AAV73924 standard; DNA; 3616 BP.
XX
AC AAV73924;
XX
DT 04-MAR-1999 (first entry)
XX
DE Human SAHH DNA #1.
XX
KW 5-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;
KW treatment; infection; cancer; autoimmune disease; detection; diagnosis;
KW gene mapping; antisense; therapy; antagonist; immunoassay; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 56..158
FT /*lag- a
FT /product= "SAHH"
XX
PN US5854023-A.
XX
PD 29-DEC-1998.
XX
PF 17-JUL-1997; 97US-0896005.
XX
PR 17-JUL-1997; 97US-0896005.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Corley NC, Hillman JL, Lal P, Shah P;
XX
DR WPI: 1999-094906/08.
DR P-PSDB; AAM90061.
XX
PT Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase -
PT for production of recombinant enzyme, useful for diagnosis,
PT treatment and prevention of cancers, infections and autoimmune
PT diseases
XX
PS Disclosure; Fig 1A-I; 40pp; English.
XX
SS This sequence encodes a human S-adenosyl-5-homocysteine hydrolase (SAHH).
CC The SAHH protein can be used to generate specific antibodies and in drug
CC screening to identify specific binding agents. Antagonists of the
CC protein are used to treat or prevent a wide range of viral, bacterial,
CC fungal, parasitic, protozoal or helminthic infections, many cancers
CC (leukemia, lymphoma or solid tumors), and many autoimmune diseases
CC (e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes
CC mellitus, multiple sclerosis etc). All these conditions may be treated by
CC expressing antisense sequences, triplex-forming agents or ribozymes
CC directed against the nucleic acid. The nucleic acid and its fragments can
CC be used as probes or primers for detecting and quantifying gene
CC expression, for diagnosis or monitoring of disease, to identify genetic
CC variations, mutations or polymorphisms, in gene mapping and as antisense
CC therapeutics. Antibodies are used directly as antagonists, indirectly to
CC deliver active agents to SAHH-expressing cells, to diagnose and monitor
CC diseases in standard immunoassays, in competitive drug screens and to
CC isolate the protein from natural sources.
XX
SQ Sequence 3616 BP; 1017 A; 782 C; 826 G; 991 T; 0 other;
Query Match 59.4%; Score 1522; DB 20; Length 3616;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1572; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 292 GGGAGAGCTGAGCAGGAGCCAGAGATGAGAGAGCCGAGAGTACTCTCATGGCCA 351
3 GGGAGAGCTGAGCAGGAGCCAGAGATGAGAGAGCCGAGAGTACTCTCATGGCCA 62

```

OY	1432	TGCGTATGCAGTAATTGGGCCCACTCCAAACACAGAATAATCATGTGACAGAGCTCCGCACTCCGG	1491
OY	1432	TGCGTATGCAGTAATTGGGCCCACTCCAAACACAGAATAATCATGTGACAGAGCTCCGCACTCCGG	1491
Db	1143	TCGTATGCAATAATTTGGCCACTCCACACAGAAAATCGATGTGACCAGCCTCCGCACTCCGG	1202
OY	1492	AGCTACACGTGGGAGGAGTAGTACGTTCTCAGGTGGACCATGTCACTCTGGCCAGATGGCAAC	1551
Db	1203	AGCTACACGTGGGAGGAGTAGTACGTTCTCAGGTGGACCATGTCACTCTGGCCAGATGGCAAC	1262
OY	1552	GAGTTGTCTCTCTGGCAGAGGGTGCTCTACTCAATTTAGAGCTGCTCCACAGTTCCCACT	1611
Db	1263	GAGTTGTCTCTCTGGCAGAGGGTGCTCTACTCAATTTAGAGCTGCTCCACAGTTCCCACT	1322
OY	1612	TGTGTCTGTCCATCACAGCCACAACAGAGCTTTGGCAGCTGATAGAACCTATATATGAC	1671
Db	1323	TGTGTCTGTCCATCACAGCCACAACAGAGCTTTGGCAGCTGATAGAACCTATATATGAC	1382
OY	1672	CCGAGGGGGGATFACACAGAGATGTGTACTTGCTGCTTAGAAGAAATGATGAATACGTTG	1731
Db	1383	CCGAGGGGGGATFACACAGAGATGTGTACTTGCTGCTTAGAAGAAATGATGAATACGTTG	1442
OY	1732	CCAGCTTGCAATCTGCCATCATTTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAA	1791
Db	1443	CCAGCTTGCAATCTGCCATCATTTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAA	1502
OY	1792	AATATCTGGGACTCAACAAAAATGGGCCATTCAAACTTAATTATACAGATGACTAATGGA	1851
Db	1503	AATATCTGGGACTCAACAAAAATGGGCCATTCAAACTTAATTATACAGATGACTAATGGA	1562
OY	1852	CCATCTACTACCAAG 1864	
Db	1563	CCATCTACTACCAAG 1575	
RESULT 5			
ID	AAF16004	AAF16004 standard; cDNA; 721 BP.	
XX	AAF16004;		
DE	13-MAR-2001	(first entry)	
XX	Human prostate cancer antigen nucleotide sequence SEQ ID NO:439.		
XX	Human; prostate cancer; prostate cancer antigen; detection; diagnosis;		
KW	neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;		
KW	vulnerability; gastrointestinal; nephrotoxic; antineoplastic; gynaecological;		
KW	antibacterial; gene therapy; neural; immune; reproductive; renal;		
KW	gastrointestinal; pulmonary; cardiovascular; proliferative disorder;		
KW	wound; infectious disease; ss.		
OS	Homo sapiens.		
XX	WO200055174-A1.		
XX	21-SEP-2000.		
XX	08-MAR-2000; 2000MO-US05988.		
PF	12-MAR-1999; 99US-0124270.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
PA	(ROSE/) ROSEN C A.		
PI	Rosen CA, Ruben SM;		
DR	WPI: 2000-587513/55.		
XX	P-PsDB; AAB56801.		
PT	Prostate cancer associated gene sequences, referred to as prostate		
PR	cancer antigens, useful for treatment, prevention, and diagnosis of		
PS	disorders such as prostate cancer		
XX	Claim 1; Page 917-918; 233bp; English.		


```

Db 241 TTTTTCATCTCATATCATATCATATCTGACAGCCACACAGGAATGCTTCATGGCTC 300
Qy 2028 TTTAGATGAATAGAGATTCAGGCTCCCTCAGCTAGTACTAGTAAGAAGATTTACTCC 2087
Db 301 TTTAGATGAATAGAGATTCAGGCTCCCTCAGCTAGTACTAGTAAGAAGATTTACTCT 360
Qy 2088 CCCAGCCCGAAGAAAGTGATTTCTCTCTTTTACCAATTTCTGGGACTTATAGTCTTAATAGG 2147
Db 361 CCCAGCCCGAAGAAAGTGATTTCTCTCTTTTACCAATTTCTGGGACTTATAGTCTTAATAGG 419
Qy 2148 TACCTATTATACAGGAATAGTACTAGTACCTCTCTGAGAACAAATCTGCAATGCTTAA 2207
Db 420 TACCTATTATACAGGAATAGTACTAGTACCTCTCTGAGAACAAATCTGCAATGCTTAA 479
Qy 2208 TCGCCTTAAAGAGCCCATTTCTTAGTGTGAATCAGTGTCTCTTCTCAGTCTTCTCAGAG 2267
Db 480 TCGCCTTAAAGAGCCCATTTCTTAGTGTGAATCAGTGTCTCTTCTCAGTCTTCTCAGAG 539
Qy 2268 AACGAGGATGTGACTACCCGCGAGGTAGTATAGATGTGGGTGTCATGTTAATTTCC 2327
Db 540 AACGAGGATGTGACTACCCGCGAGGTAGTATAGATGTGGGTGTCATGTTAATTTCC 599
Qy 2328 CTAGAGAGTCCAGCCCTGTTCTGCTGCTAAAGTGTGATGTCACAGTACAGATGCT 2387
Db 600 CTAGAGAGTCCAGCCCTGTTCTGCTGCTAAAGTGTGATGTCACAGTACAGATGCT 659
Qy 2388 ATATATAGCATGCTTGTATAGATCAGAGAGCCCATCTGATTTATATAGCTTCC 2447
Db 660 ATATATAGCATGCTTGTATAGATCAGAGAGCCCATCTGATTTATATAGCTTCC 719
Qy 2448 TCCACTCCACCAAGACTTGTCTATTTTTCAGTTTAACTAGACTACACTCTATT 2503
Db 720 TCCACTCCACCAAGACTTGTCTATTTTTCAGTTTAACTAGACTACACTCTATT 775

RESULT 7
ABV44141/C
ID ABV44141 standard; cDNA; 504 BP.
AC ABV44141;
XX
XX 16-SEP-2002 (first entry)
DE Human prostate expression marker cDNA 44132.
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
OS Homo sapiens.
XX
XX MO200160860-A2.
PN
XX
XX 23-NOV-2001.
PD
XX
XX 20-FEB-2001; 2001WO-US05171.
PF
XX
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer. Useful
XX for detecting presence of prostate cancer, stage of prostate cancer
XX

```

```

XX
XX Claim 1, Page 8770; 11750pp; English.
PS
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 504 BP; 117 A; 150 C; 116 G; 121 T; 0 other;
XX
XX Query Match 16.5%; Score 422; DB 23; Length 504;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-151;
XX Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 706 GTACACACATCACAGCCACAGACGGGTGATTGATGACACACTGTGCCCCGGGGGCTG 765
XX Db 502 GTACACACATCACAGCCACAGACGGGTGATTGATGACACACTGTGCCCCGGGGGCTG 443
XX Qy 766 AGTCCCGCTGCTGCTTGTATCACTCTACTCACTCAGATGAAATGAACTGACGACTGG 825
XX Db 442 AGTCCCGCTGCTGCTTGTATCACTCTACTCACTCAGATGAAATGAACTGACGACTGG 383
XX Qy 826 CTAGAGGTGAGTGTGAGTGTGCTGCTGGAAGGGGAGTCAAGATGACTTCTGCTGT 885
XX Db 382 CTAGAGGTGAGTGTGAGTGTGCTGCTGGAAGGGGAGTCAAGATGACTTCTGCTGT 323
XX Qy 886 GTATTTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTGGATGATGGG 945
XX Db 322 GTATTTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCTGGATGATGGG 263
XX Qy 946 GAACTTTAACCCACTGGGTATTATAGAAAGTATCCAAAGCTTTTAAAGATCCGAGGCA 1005
XX Db 262 GAACTTTAACCCACTGGGTATTATAGAAAGTATCCAAAGCTTTTAAAGATCCGAGGCA 203
XX Qy 1006 TTGTGGAAGAGAGCTGACTGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC 1065
XX Db 202 TTGTGGAAGAGAGCTGACTGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC 143
XX Qy 1066 TCTGTGTCCGGCCATGAGTCAATGATCTGTACCAAAAGATTTGATTAATCTGT 1125
XX Db 142 TCTGTGTCCGGCCATGAGTCAATGATCTGTACCAAAAGATTTGATTAATCTGT 83
XX Qy 1126 AC 1127
XX Db 82 AC 81

RESULT 8
AAZ80766/C
ID AAZ80766 standard; cDNA; 636 BP.
XX
XX AAZ80766;
AC
XX
XX 07-APR-2000 (first entry)
DE Human colon cancer cell line SW480 cDNA clone SEQ ID NO:850.
XX
XX Human colon cancer cell line SW480 cDNA clone SEQ ID NO:850.
XX
XX Human: gene expression product; diagnosis; tumour; colon cancer;
XX colorectal adenocarcinoma; cell line SW480; cell proliferation;
XX cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
XX hyperplasia; ds.
XX

```

OS Homo sapiens.
 XX MO964576-A2.
 XX 16-DEC-1999.
 XX 09-JUN-1999; 99MO-IB01062.
 XX 10-JUN-1998; 98US-0088801.
 XX (FARB) BAYER CORP.
 PI Endege WO, Steinhann KE, Aetle JR, Burgess CC, Bushnell SE;
 PI Carroll E, Catino TJ, Dertl A, Ford DM, Lewis ME, Monahan JE;
 PI Schlegel R;
 XX MPI; 2000-087220/07.
 DR Novel nucleic acids, used to develop products for the diagnosis and
 PT treatment of disorders involving unwanted cell proliferation,
 PT particularly cancers, especially colon cancer
 PS Claim 15; Page 469; 469pp; English.
 XX AA279917 to AA280766 represent double stranded cDNA clones isolated from
 CC the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
 CC cDNA clones can be used to generate antisense oligonucleotides which
 CC can be used for antisense therapy. Methods and products from the present
 CC invention can be used for identifying and/or classifying cancerous cells
 CC present in a human tumour, particularly in solid tumours, e.g.
 CC carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
 CC can be used for developing agents for the diagnosis and treatment of
 CC disorders involving unwanted cell proliferation, such as neoplasia,
 CC dysplasia or hyperplasia.
 SQ Sequence 636 BP; 156 A; 166 C; 137 G; 147 T; 30 other:
 Query Match 14.4%; Score 370; DB 21; Length 636;
 Best Local Similarity 100.0%; Pred. No. 1.2e-131;
 Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 756 CTGGGGGCTCAGCCCGCTGCTGCTGAACATCTACTCAACTCAGAAAGTAGCT 815
 DB 370 CTGGGGGCTCAGCCCGCTGCTGCTGAACATCTACTCAACTCAGAAAGTAGCT 311
 QY 816 GCACGACTGCTGAGGCTGAGTTCAGTTCGCTTGAAGGGGAGTCAAGATGAC 875
 DB 310 GCACGACTGCTGAGGCTGAGTTCAGTTCGCTTGAAGGGGAGTCAAGATGAC 251
 QY 876 TTCTGGTGTGTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGATCTG 935
 DB 250 TTCTGGTGTGTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGATCTG 191
 QY 936 GATGATGGGGAGACTTAACCACTGCTTATAGAGTATCCAAAGCTTTAAGAAG 995
 DB 190 GATGATGGGGAGACTTAACCACTGCTTATAGAGTATCCAAAGCTTTAAGAAG 131
 QY 996 ATCCGAGGACTTGTGAAGAGAGGCTGAGTTCACAGGCTGATACGCTCCAAA 1055
 DB 130 ATCCGAGGACTTGTGAAGAGAGGCTGAGTTCACAGGCTGATACGCTCCAAA 71
 QY 1056 GCTGGAGGCTCTGTGTCGCGCCATGAGATGATGTTCTTTACCAAGAAATT 1115
 DB 70 GCTGGAGGCTCTGTGTCGCGCCATGAGATGATGTTCTTTACCAAGAAATT 11
 QY 1116 GATTAAGTGT 1125
 DB 10 GATTAAGTGT 1

XX ABN94533;
 AC 13-AUG-2002 (first entry)
 XX 13-AUG-2002 (first entry)
 DT Gene #1031 used to diagnose liver cancer.
 DE Gene #1031 used to diagnose liver cancer.
 XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatocytic;
 KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX Homo sapiens.
 OS Homo sapiens.
 XX MO200229103-A2.
 XX 11-APR-2002.
 PD 02-OCT-2001; 2001MO-US30589.
 XX 02-OCT-2001; 2001MO-US30589.
 XX 02-OCT-2001; 2000US-237054P.
 PR (GENE-) GENE LOGIC INC.
 PA Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 PI MPI; 2002-426119/45.
 DR Diagnosing and detecting the progression of liver cancer.
 XX hepatocellular carcinoma or metastatic liver tumor in a patient.
 PT Involves detecting the level of expression of two or more genes in a
 PT liver tissue sample
 PS Claim 1; SEQ ID NO 1031; 298pp; English.
 XX The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytosolic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 SQ Sequence 376 BP; 125 A; 78 C; 81 G; 92 T; 0 other:
 Query Match 12.9%; Score 330; DB 24; Length 376;
 Best Local Similarity 100.0%; Pred. No. 2.4e-116;
 Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2174 TACCTTCTGTGTGAACAACTGCAATGTCTTAATGCTTAAAGAGCCATTCTTG 2233
 DB 376 TACCTTCTGTGTGAACAACTGCAATGTCTTAATGCTTAAAGAGCCATTCTTG 317
 QY 2234 CTGCTGAATCAGTGTCTTCACTTCTTCAAGAGAGGAGTGTACTCCGGGCG 2293
 DB 316 CTGCTGAATCAGTGTCTTCACTTCTTCAAGAGAGGAGTGTACTCCGGGCG 257
 QY 2294 GTAGGTAGATGTGGGTGTCATGTTAATTTCCCTTGAAGTTCCAGCCCTGTTTCT 2353
 DB 256 GTAGGTAGATGTGGGTGTCATGTTAATTTCCCTTGAAGTTCCAGCCCTGTTTCT 197
 QY 2354 GCGTAAAGGTGTATGTCACGATGATGATTAATGACATGCTGTTGAAGATCA 2413
 DB 196 GCGTAAAGGTGTATGTCACGATGATGATTAATGACATGCTGTTGAAGATCA 137
 QY 2414 GGAGGCCACTGTGATTTATAGTATAGCCCTTCCACACTCCACAGACTGTCATTT 2473

RESULT 9
 ABN94533/c
 ID ABN94533 standard; DNA; 376 BP.

XX Claim 1; Fig 1; 299bp; English.

CC The present invention describes an isolated polypeptide comprising an
 CC immunogenic portion of an ovarian carcinoma protein (or its variants).
 CC Ovarian carcinoma proteins, and polynucleotides encoding them, have
 CC cytostatic activity and can be used in gene therapy and vaccines.
 CC Ovarian carcinoma polypeptides, nucleic acids, antibodies and vaccines
 CC are useful for the prevention, diagnosis and treatment of cancer.
 CC Preferably ovarian cancer. AA69691 to AA70077 and AB12552 to AB12557
 CC represent human ovarian carcinoma polynucleotides and proteins used in
 CC the exemplification of the present invention.

CC Sequence 292 BP; 98 A; 68 C; 60 G; 66 T; 0 other;

XX Query Match 9.2%; Score 237; DB 21; Length 292;

XX Best Local Similarity 100.0%; Pred. No. 7.8e-81; Mismatches 0; Gaps 0;

XX Matches 237; Conservative 0; Indels 0; Indels 0; Gaps 0;

DB 2267 GAACGAGGATGTACTACCCGCGAGGTAGTATGTGGTGGTGCATGTAATTC 2326

DB 276 GAACGAGGATGTACTACCCGCGAGGTAGTATGTGGTGGTGCATGTAATTC 217

QY 2327 CCTTGAAGTTCGAAGCCCTGTTCCGTAAGGTGGTATGTCAGTTCAGATGTG 2386

DB 216 CCTTGAAGTTCGAAGCCCTGTTCCGTAAGGTGGTATGTCAGTTCAGATGTG 157

QY 2387 TATAATGAGCATGGCTGTTTAAGATCAGAGGCCCACTGGATTATATAGCCCTTC 2446

DB 156 TATAATGAGCATGGCTGTTTAAGATCAGAGGCCCACTGGATTATATAGCCCTTC 97

QY 2447 CTCACATCCCAACAGACTGCTCATTTTTCAGATTGTTTAACTAGACTACTATT 2503

DB 96 CTCACATCCCAACAGACTGCTCATTTTTCAGATTGTTTAACTAGACTACTATT 40

RESULT 12

ABN72624/c

ABN72624 standard; DNA; 292 BP.

AC ABN72624;

DT 02-JUL-2002 (first entry)

DE Ovarian carcinoma antigen polynucleotide #40.

DE Human; immunostimulant; cytostatic; cancer; ovarian carcinoma; ds.

OS Homo sapiens.

PN WO200206317-A2.

PD 24-JAN-2002.

PF 17-JUL-2001; 2001WO-US22635.

PR 17-JUL-2000; 2000US-0617747.

PR 10-AUG-2000; 2000US-0636801.

PR 20-SEP-2000; 2000US-0667857.

PR 04-APR-2001; 2001US-083271.

PR 18-JUN-2001; 2001US-0884441.

XX (CORI-) CORIXA CORP.

PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;

PI Reed SG, Vedvick TS, Carter D, Hill P, Albond E;

DR WPI; 2002-164781/21.

XX Polypeptides comprising an immunogenic portion of an ovarian carcinoma

PT protein or its variants, useful for stimulating an immune response in a

PT patient and treating ovarian cancer.

PS Example 1; Page 104; 408bp; English.

CC This invention relates to polypeptides comprising an immunogenic
 CC portion of an ovarian carcinoma protein which acts as an
 CC immunostimulant and is cytostatic. The polypeptides, polynucleotides,
 CC antibodies, fusion proteins, T cell populations and antigen presenting
 CC cells that express the polypeptides are useful for stimulating an
 CC immune response in a patient and treating ovarian cancer. This
 CC sequence represents DNA related to the invention.

CC Sequence 292 BP; 98 A; 68 C; 60 G; 66 T; 0 other;

XX Query Match 9.2%; Score 237; DB 24; Length 292;

XX Best Local Similarity 100.0%; Pred. No. 7.8e-81; Mismatches 0; Gaps 0;

XX Matches 237; Conservative 0; Indels 0; Indels 0; Gaps 0;

QY 2267 GAACGAGGATGTACTACCCGCGAGGTAGTATGTGGTGGTGCATGTAATTC 2326

DB 276 GAACGAGGATGTACTACCCGCGAGGTAGTATGTGGTGGTGCATGTAATTC 217

QY 2327 CCTTGAAGTTCGAAGCCCTGTTCCGTAAGGTGGTATGTCAGTTCAGATGTG 2386

DB 216 CCTTGAAGTTCGAAGCCCTGTTCCGTAAGGTGGTATGTCAGTTCAGATGTG 157

QY 2387 TATAATGAGCATGGCTGTTTAAGATCAGAGGCCCACTGGATTATATAGCCCTTC 2446

DB 156 TATAATGAGCATGGCTGTTTAAGATCAGAGGCCCACTGGATTATATAGCCCTTC 97

QY 2447 CTCACATCCCAACAGACTGCTCATTTTTCAGATTGTTTAACTAGACTACTATT 2503

DB 96 CTCACATCCCAACAGACTGCTCATTTTTCAGATTGTTTAACTAGACTACTATT 40

RESULT 13

ABL79045

ABL79045 standard; CDNA; 553 BP.

AC ABL79045;

DT 17-MAY-2002 (first entry)

DE Human ovarian cancer related CDNA clone SEQ ID NO:2023.

DE Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

OS Homo sapiens.

PN WO200192581-A2.

PD 06-DEC-2001.

PF 29-MAY-2001; 2001WO-US17756.

PR 26-MAY-2000; 2000US-207484P.

XX (CORI-) CORIXA CORP.

PI Algate PA, Harlocker SL, Jones R;

PI WPI; 2002-122075/16.

XX Composition for therapy and diagnosis of ovarian cancer comprising

PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding

PT polypeptide, antibody specific to polypeptide or T cell expressing

PT polypeptide.

PS Claim 1; SEQ ID 2023; 489bp; English.

XX The present invention describes a composition (I) comprising: carriers

CC and immunostimulants; and a polypeptide (II) of a ovarian tumour

CC polypeptide encoded by a polynucleotide (III) having a CDNA sequence

CC (SI) from the 10912 nucleotide sequences as given in ABL77023 to

CC ABL87934, (III) encoding (II) having a sequence (52), a T cell

KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX
 OS Homo sapiens.
 XX
 PN WO200229103-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 02-OCT-2001; 2001WO-US30589.
 XX
 PR 02-OCT-2000; 2000US-237054P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 XX WPI; 2002-426119/45.
 DR
 XX
 PT Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample
 XX
 PS Claim 1; SEQ ID NO 346; 298bp; English.
 XX
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 384 BP; 95 A; 77 C; 113 G; 99 T; 0 other;

Query Match 6.6%; Score 168; DB 24; Length 384;
 Best Local Similarity 100.0%; Pred. No. 1.6e-54;
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1065 CTCGTGTCGCGCCGANGACGTCATGATCTCTGTACCAACAGAAAGTTGATAACTTG 1124
 |||||||
 DB 73 CTCGTGTCGCGCCGANGACGTCATGATCTCTGTACCAACAGAAAGTTGATAACTTG 132
 OY 1125 TACTGCTGCCGAGATCATTTTGGATGGCCCTGAAGAGACACAGATGTGATTTGGT 1184
 |||||||
 DB 133 TACTGCTGCCGAGATCATTTTGGATGGCCCTGAAGAGACACAGATGTGATTTGGT 192
 OY 1185 GGGAAACAAGTGTGTGTGTGGCTATGTGTAGGTAGGCAAGGGCTGC 1232
 |||||||
 DB 193 GGGAAACAAGTGTGTGTGTGGCTATGTGTAGGTAGGCAAGGGCTGC 240

Search completed: April 20, 2003, 19:20:50
 Job time : 561 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 11:15:10 ; Search time 6526 Seconds
(without alignments)
11429.741 Million cell updates/sec

Title: US-09-782-051-1

Sequence: 1 ggcgcggcagctcgagct.....aaaaaaaaaaaaaaaaaaaa 2563

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.ov:*
22: em.pat:*
23: em.ph:*
24: em.pl:*
25: em.ro:*
26: em.sts:*
27: em.un:*
28: em.vl:*
29: em.yl:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pin:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2563	100.0	2563	6 AX029176	AX029176 Sequence
2	2563	100.0	2577	9 AF315687	AF315687 Homo sapi
3	2531	98.8	2583	9 BC010681	BC010681 Homo sapi
4	2525.2	98.5	2552	9 BC016942	BC016942 Homo sapi
5	2495.6	97.4	2526	9 BC007576	BC007576 Homo sapi
6	2218.6	86.6	2258	9 HS082761	HS082761 Homo sapi
7	2161.2	84.3	2510	9 HSM800298	AL049954 Homo sapi
8	1525.4	59.5	1772	10 BC018218	BC018218 Mus muscu
9	1333.8	52.0	1369	9 AF035319	AF035319 Homo sapi
10	1288.8	50.3	1323	9 BC003631	BC003631 Homo sapi
11	1001.6	39.1	2030	9 BC008349	BC008349 Homo sapi
12	1001.6	39.1	5052	9 BC024325	BC024325 Homo sapi
13	998	38.9	5025	9 AB020635	AB020635 Homo sapi
14	982.2	38.3	1987	9 AK025372	AK025372 Homo sapi
15	806.2	31.5	2226	9 AR068569	AR068569 Sequence
16	806.2	31.5	2226	6 AX069339	AX069339 Sequence
17	703.2	27.4	81378	9 AL772411	AL772411 Human DNA
18	703.2	27.4	183584	30 AL360171	AL360171 Human DNA
19	655.2	25.6	1792	3 AY113501	AY113501 Drosophila
20	488.2	19.0	40320	2 AC015387	AC015387 Drosophila
21	488.2	19.0	172372	3 AC010110	AC010110 Drosophila
22	421.6	16.4	52423	2 AC020395	AC020395 Drosophila
23	421.6	16.4	80423	3 DROABDB	L07835 Drosophila
24	421.6	16.4	175335	3 AC091636	AC091636 Drosophila
25	421.6	16.4	223098	3 AE003715	AE003715 Drosophila
26	421.6	16.4	338234	3 DM031961	U31961 Drosophila
27	414	16.2	2057	10 M055AH	L32836 Mus musculu
28	412.4	16.1	2067	10 BC015304	BC015304 Mus muscu
29	412.4	16.1	190612	2 AC129093	AC129093 Mus muscu
30	411.6	16.1	5050	3 DMBX200	X13168 Drosophila
31	407.4	15.9	1302	6 AX458351	AX458351 Sequence
32	404.6	15.8	2029	10 RATVHHA	M15185 Rat S-aden
33	404	15.8	2084	9 HDMAHCY2	M61832 Human S-ade
34	404	15.8	2175	9 BC010018	BC010018 Homo sapi
35	404	15.8	2183	9 BC011606	BC011606 Homo sapi
36	404	15.8	2355	9 AK097610	AK097610 Homo sapi
37	403.6	15.7	2088	5 XLA7835	AB007835 Xenopus L
38	400.4	15.6	2129	5 XELAH	L35559 Xenopus lae
39	377.8	14.7	1603	3 AY102668	AY102668 Drosophila
40	376.6	14.7	1564	3 DMAHCYGEN	X95636 D.melanog
41	365	14.2	376	6 AX408384	AX408384 Sequence
42	362	14.1	1617	3 AF080546	AF080546 Anopheles
43	357	13.6	1631	3 LEISADSH	M76556 L.donorv
44	348.4	13.6	1315	3 DDIAHHA	M19937 Slime mold
45	337.6	13.2	173839	2 AC118421	AC118421 Rattus no

ALIGNMENTS

RESULT 1
AX029176
LOCUS: AX029176 2563 bp DNA linear PAT 16-SEP-2000
DEFINITION: Sequence 1 from Patent #09814562.
ACCESSION: AX029176
VERSION: AX029176.1 GI:10190060
KEYWORDS:
SOURCE: human.
ORGANISM: Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE: Hart,D.N.
AUTHORS: 1 (bases 1 to 2563)
TITLE: Enzyme having s-adenosyl-1-homocysteine hydrolase (ahcy) type activity

JOURNAL Patent: WO 9814562-A 1 09-APR-1998;
HART DEREK NIGEL JOHN (NZ)

FEATURES Location/Qualifiers
Source 1. 2563
/organism="Homo sapiens"
/db_xref="taxon:9606"

CDS
/note="Open reading frame extends without a stop codon for the full 5' nucleotide sequence. The initiation codon has yet to be identified."
/codon_start=-1
/protein_id="CAC09285.1"
/db_xref="GI:10190061"

/translation="RGQVGRSCCFPFSCGHRCPALGCTDKAMATAPKPTQLODA
GAGRRVDVSEBAPARAPGEPERGGGSAAGRGAGGMSMPDAPLPVGEELKQ
AAEIEDAEKYSFPAATVAKPKQIOFADMQEPTKPTKGRSLRSISQSSLDYS
SAASYSDSDSEVPREKQOTNSKSNCEVKNKQAEFRGEIEIIONDSALISLR
KRAQGEPLAGAKIVGCTHTTAQAVILETICALGACRNSACIYSEONVAALAE
ACVAVAMKGESEDDPFMCIDRCVNMGNMANNMIDGDLTHVYKYPVPEKIRG
IYBESYGYHRLTOLSKAGKLCYPAAMVNDVYTKQKDNLYCCESILDGLKRTDYA
FGKQVVCGYGVGKCCALALAGALVITTEIDPLCALQACDREVRVNLVEYTDV
VDVITCTGNKRVNREHLDNRKNSCIYCNMGHSNTEIDVSLTPELTWRVRSQVD
HVIWPGKRVLLAEGRLLNLSCSTVPEFLSIATQALALIELYNAPEGRYKQDYY
LLEPKMDEVYASLHLPSEFHAHLTEITDQAKYLGKLNKNGPEKPYRYR"

BASE COUNT 646 a 604 c 677 g 636 t

ORIGIN

Query Match 100.0%; Score 2563; DB 6; Length 2563;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGGGGGGGAGTGGAGCTGCGAGCTGCTCTGTTCTCTTGGCCACCTCGT 60
1 GGGGGGGGAGTGGAGCTGCGAGCTGCTCTGTTCTCTTGGCCACCTCGT 60
61 GTCGGGCTGCTGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
61 GTCGGGCTGCTGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
121 GCGAGCTGAG 180
121 GCGAGCTGAG 180
181 GGGGAGGCGGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
181 GGGGAGGCGGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
181 GGGGAGGCGGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
241 GGGGAGGCGGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
241 GGGGAGGCGGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
301 TGAAGCAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
301 TGAAGCAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
301 TGAAGCAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
361 AGGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
361 AGGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
421 CCAAAATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
421 CCAAAATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
481 GTTCAGTGCATCTTACAGATAGCTCTGATGATGAGAGAGAGAGAGAGAG 540
481 GTTCAGTGCATCTTACAGATAGCTCTGATGATGAGAGAGAGAGAGAGAG 540
541 AAACCACTCCAG 600
541 AAACCACTCCAG 600
601 GAGCCGAGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
601 GAGCCGAGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

601 GAGCCGAGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
661 GTGCTAG 720
661 GTGCTAG 720
721 CCCAG 780
721 CCCAG 780
781 CTTGTAACTCTACTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
781 CTTGTAACTCTACTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
841 CAGTGTCCGTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
841 CAGTGTCCGTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
901 TGAACATGATGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
901 TGAACATGATGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
961 GGGTTTAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
961 GGGTTTAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
1021 TGACTGCTGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1021 TGACTGCTGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1081 TGAACATGATGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
1081 TGAACATGATGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
1141 CCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1141 CCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1201 TGTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
1201 TGTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
1261 TGTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1261 TGTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1321 GGGTGTAAAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1321 GGGTGTAAAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1381 ATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
1381 ATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
1441 ATATGGGCGAG 1500
1441 ATATGGGCGAG 1500
1501 GGGAGCGAG 1560
1501 GGGAGCGAG 1560
1561 TCTGTGAG 1620
1561 TCTGTGAG 1620
1621 CCATCAG 1680
1621 CCATCAG 1680
1681 GATACAG 1740
1681 GATACAG 1740

QY	361	AGGGCCCAAGAAACCAATCCAGTTGGCTGTATGACATGACGAGGATCCACCAATTTCCCA	420
Db	475	AGGGGCCCAAGAAACCAATCCAGTTGGCTGTATGACATGACGAGGATCCACCAATTTCCCA	534
QY	421	CCAAACATGAGCCGAAGATCTTTGTCTGCTGCATCTCACAGTCCCTCAGTACGAGCTACA	480
Db	535	CCAAACATGAGCCGAAGATCTTTGTCTGCTGCATCTCACAGTCCCTCAGTACGAGCTACA	594
QY	481	GTTCAGCTGCATCCTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGAGAACGAC	540
Db	595	GTTCAGCTGCATCCTACACAGATAGCTCTGATGATGAGTTTCTCCCGAGAGAACGAC	654
QY	541	AAACCAACTCCAAAGGCGACGACCAATTTCTGTGTGAAGAACATCAAGCGAGAAATTGG	600
Db	655	AAACCAACTCCAAAGGCGACGACCAATTTCTGTGTGAAGAACATCAAGCGAGAAATTGG	714
QY	601	GAGCCGGGAGATGAGATTGCGAGACGACACATGCTGCTGATTTACTACAGGAAC	660
Db	715	GAGCCGGGAGATGAGATTGCGAGACGACACATGCTGCTGATTTACTACAGGAAC	774
QY	661	GTGCTCAGGGGGAGAAAGCCCTTGGCTGTGCTAAATAGTGGCTGTACACATGACAG	720
Db	775	GTGCTCAGGGGGAGAAAGCCCTTGGCTGTGCTAAATAGTGGCTGTACACATGACAG	834
QY	721	CCCAAGACAGGGGTGTATTGAGACACTCTGTGCTCCCTGGGGGCTCAGTCCGCTGGTGG	780
Db	835	CCCAAGACAGGGGTGTATTGAGACACTCTGTGCTCCCTGGGGGCTCAGTCCGCTGGTGG	894
QY	781	CTTGTACATCTACTCAACTCAGAAATAGTACTGACAGACTGGGTGAGGGCTGGAGTTG	840
Db	895	CTTGTACATCTACTCAACTCAGAAATAGTACTGACAGACTGGGTGAGGGCTGGAGTTG	954
QY	841	CAGTGTTCGCTTGGAGGGCGGAGCTCAGAGATGACTTCTGGTGGTATTGACCGCTGTG	900
Db	955	CAGTGTTCGCTTGGAGGGCGGAGCTCAGAGATGACTTCTGGTGGTATTGACCGCTGTG	1011
QY	901	TGAACATGAGATGGGTGGCAGGCCAACATATCTGTGATGATGGGGGAGACTTAACCACT	960
Db	1015	TGAACATGAGATGGGTGGCAGGCCAACATATCTGTGATGATGGGGGAGACTTAACCACT	1077
QY	961	GGGTTTAAAGAAATATCCAAACGCTGTTTAAAGAAATCCAGGCACTTGTGGAAAGAGGG	1021
Db	1075	GGGTTTAAAGAAATATCCAAACGCTGTTTAAAGAAATCCAGGCACTTGTGGAAAGAGGG	1133
QY	1021	TGACTGCTGTTCACAGGCTGTATCAGCTCCCAAAGCTGGGAAGCTCTGTGTTCCGGCA	1081
Db	1135	TGACTGCTGTTCACAGGCTGTATCAGCTCCCAAAGCTGGGAAGCTCTGTGTTCCGGCA	1193
QY	1081	TGAACGTCATGATTTCTGTTACCAACAGAGTTTGATTAATCTGTACTGTGCCGAAT	1144
Db	1195	TGAACGTCATGATTTCTGTTACCAACAGAGTTTGATTAATCTGTACTGTGCCGAAT	1257
QY	1141	CCATTGAGATGGCTGAAGAGACACAGATGATGATTTGGTGGGAAACAAGTGGTG	1201
Db	1255	CCATTGAGATGGCTGAAGAGACACAGATGATGATTTGGTGGGAAACAAGTGGTG	1311
QY	1201	TGTTGGCTATGTTGAGAGTGAAGCAAGGGCTGTGTCTGTCTCAAAAGCTTTGGAGCA	1261
Db	1315	TGTTGGCTATGTTGAGAGTGAAGCAAGGGCTGTGTCTGTCTCAAAAGCTTTGGAGCA	1377
QY	1261	TTTGTCTCATTTACCGAATGACCCCATCTGTCTCTGACAGGCTCATGAGATGGTTCA	1321
Db	1375	TTTGTCTCATTTACCGAATGACCCCATCTGTCTCTGACAGGCTCATGAGATGGTTCA	1433
QY	1321	GGGTTGAAGAGCTTAAAGAGTCAATCCGCAATCGATGTCGTAATTAATTTGCACAGAA	1381
Db	1435	GGGTTGAAGAGCTTAAAGAGTCAATCCGCAATCGATGTCGTAATTAATTTGCACAGAA	1499
QY	1381	ATAAGATATGATGACACAGGAGACACTTGGATGCGATGAAAAACAGTTTGTATTCGTGCA	1441
Db	1495	ATAAGATATGATGACACAGGAGACACTTGGATGCGATGAAAAACAGTTTGTATTCGTGCA	1551
QY	1441	ATATGGCCACTCCAACACAGAAATGATGTGACCAAGCCTCCGCACTCCGGAGGTGACGT	1501

Db	1555	ATATGGGGCACTCCAAACACAGAAATGATGTGACACACCTCCGGACCTGGAGCTGACGT	1614
Qy	1501	GGGAGGAGTACCTTTCTCAGGTGGAGCAATGTCAATCTGGCAGATGGCAAAAGATTTGTC	1560
Db	1615	GGGAGGAGTACCTTTCTCAGGTGGAGCAATGTCAATCTGGCAGATGGCAAAAGATTTGTC	1674
Qy	1561	TCTTGCGAGAGGGTGGTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTGTCTGT	1620
Db	1675	TCTTGCGAGAGGGTGGTCTACTCAATTTGAGCTGCTCCACAGTTCACCTTTGTCTGT	1734
Qy	1621	CCATCACAGCCACAAACACAGGCTTTGGCACTGATAGAACTCTATTAATGCAACCCAGGGGC	1688
Db	1735	CCATCACAGCCACAAACACAGGCTTTGGCACTGATAGAACTCTATTAATGCAACCCAGGGGC	1794
Qy	1681	GATPACAGCAGAGTGTGTACTTCTGCTTCTTAAGAAAAAGATGATAGCTTCCAGCTTGC	1740
Db	1795	GATPACAGCAGAGTGTGTACTTCTGCTTCTTAAGAAAAAGATGATAGCTTCCAGCTTGC	1855
Qy	1741	ATCTGCCATCATTTTGATGTGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	1800
Db	1855	ATCTGCCATCATTTTGATGTGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	1914
Qy	1801	GACTCAACAAAAATGGGGCCATTCAAACTTAATTACAGATACTAATGACCATCTAC	1866
Db	1915	GACTCAACAAAAATGGGGCCATTCAAACTTAATTACAGATACTAATGACCATCTAC	1974
Qy	1861	CAAGGACACAGTCCACGTGAAACACACACTCTAAAGAAATATTTTAAAGATACCTTAT	1920
Db	1975	CAAGGACACAGTCCACGTGAAACACACACTCTAAAGAAATATTTTAAAGATACCTTAT	2033
Qy	1921	TTTTCTTCTTACTCTCTTCTCTTGATTTTTTTTCTTAAATTTCAATCTGTGTTTTTCATC	1986
Db	2035	TTTTCTTCTTACTCTCTTCTCTTGATTTTTTTTCTTAAATTTCAATCTGTGTTTTTCATC	2094
Qy	1981	TCATTTCCAAAGTTCTTGACAGACACACAGAAACTGTGTTATGCTCTTTAGATGAATA	2040
Db	2095	TCATTTCCAAAGTTCTTGACAGACACACAGAAACTGTGTTATGCTCTTTAGATGAATA	2154
Qy	2041	GAAGTTCAGGGTCCCTCAGCTCTAGTCACTTAAGAAAGATTTTACTCCCCAGCCAGAAA	2100
Db	2155	GAAGTTCAGGGTCCCTCAGCTCTAGTCACTTAAGAAAGATTTTACTCCCCAGCCAGAAA	2214
Qy	2101	GGTGATTTCTTCTTTACCAATTTCTGGGACCTTAGCTTAATTAGTACTCTATTAAACA	2160
Db	2215	GGTGATTTCTTCTTTACCAATTTCTGGGACCTTAGCTTAATTAGTACTCTATTAAACA	2274
Qy	2161	GGAAATGCTAAGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAATCGCCTTAAAGA	2220
Db	2275	GGAAATGCTAAGTACCTTCTCTGTGGAACAATCTGCAATGTCTAAATCGCCTTAAAGA	2334
Qy	2221	GGCCATTTCTTAGCTGCTGAAACAGAGCGCTTCACTTCTTCAGAAAGAGGAGATGCT	2280
Db	2335	GGCCATTTCTTAGCTGCTGAAACAGAGCGCTTCACTTCTTCAGAAAGAGGAGATGCT	2394
Qy	2281	ACCTACCCGCGAGGTAGGTAGATGTGGGTGGTGCATGTTAATTTCCCTTGAAGTTCCA	2340
Db	2395	ACCTACCCGCGAGGTAGGTAGATGTGGGTGGTGCATGTTAATTTCCCTTGAAGTTCCA	2454
Qy	2341	AGCCCTGTTTCTCGCTAAAGGTGTATGTCCAGTTCCAGAGATGTGTAAATGACATGG	2400
Db	2455	AGCCCTGTTTCTCGCTAAAGGTGTATGTCCAGTTCCAGAGATGTGTAAATGACATGG	2514
Qy	2401	CTTGTTAAGATCAGAGGGCCACTTGATTTAATAGTATACCCCTTCCACATCCACACA	2466
Db	2515	CTTGTTAAGATCAGAGGGCCACTTGATTTAATAGTATACCCCTTCCACATCCACACA	2574
Qy	2461	GACTTGCTCATTTTGTAGTTTTTAACTAGACTACACTATTTGAGTTAATTTGTCC	2520
Db	2575	GACTTGCTCATTTTGTAGTTTTTAACTAGACTACACTATTTGAGTTAATTTGTCC	2634
Qy	2521	TCTAGATTTATTTCTGTTGTCCAAAAAATTTTTTTTTTAAAAA 2563	


```

DB 1103 TGAACGTCATGATTCGTGTAACCAACAGAAAGTTGATTAACCTGTACGTCGCGAGAT 1162
DB 1141 CCATTTTGGATGCGCTGGAAGACACACATGATGATTTGGTGGGAAACAAAGTGGGG 1200
DB 1163 CCATTTTGGATGCGCTGGAAGACACACATGATGATTTGGTGGGAAACAAAGTGGGG 1222
DB 1201 TGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1223 TGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1282
DB 1261 TGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1283 TGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1342
DB 1321 GGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1343 GGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1402
DB 1381 ATAGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1403 ATAGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1462
DB 1441 ATAGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1463 ATAGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1522
DB 1501 GGGAGCAGTACGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1523 GGGAGCAGTACGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1582
DB 1561 TCGTGGAGAGGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1583 TCGTGGAGAGGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1642
DB 1621 CCATGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1680
DB 1643 CCATGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1702
DB 1681 GATGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1740
DB 1703 GATGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1762
DB 1741 ATCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
DB 1763 ATCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1822
DB 1801 GACTCAACAAATGAGGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 1823 GACTCAACAAATGAGGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1882
DB 1861 CAAGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
DB 1883 CAAGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1942
DB 1921 TTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
DB 1943 TTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2002
DB 1981 TCAATTAACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
DB 2003 TCAATTAACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2062
DB 2041 GAAGTTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
DB 2063 GAAGTTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2122
DB 2101 GGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
DB 2123 GGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2181
DB 2161 GGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220

```

```

DB 2182 GGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2241
DB 2221 GGCATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
DB 2242 GGCATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2301
DB 2281 ACCATCCCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
DB 2302 ACCATCCCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2361
DB 2341 AGCCCTGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
DB 2362 AGCCCTGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2421
DB 2401 CTTGTTAAGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2460
DB 2422 CTTGTTAAGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2481
DB 2461 GACTGCTCATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
DB 2482 GACTGCTCATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2540
DB 2521 TCTAGATTTATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2563
DB 2541 TCTAGATTTATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2583

```

```

RESULT 4
LOCUS BC016942 2552 bp mRNA linear PRI 09-NOV-2001
DEFINITION Homo sapiens, S-adenosylhomocysteine-hydrolase-like 1, clone
ACCESSION BC016942
VERSION MG:21453 IMAGE:3450568, mRNA, complete cds.
KEYWORDS BC016942.1 GI:16877386
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2552)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.sngc.stanford.edu
Contact: (Dickson, Mark) mcdick@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Series: IRAC Plate: 20 Row: a Column: 15 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES

```

source
1..2552
/organism="Homo sapiens"
/db_xref="locustid:10768"
/db_xref="taxon:9606"
/clone="MGC:21453 IMAGE:3450568"
/tissue_type="Placenta, chorioncarcinoma"
/clone_lib="NIH_MGC_10"

```

/lab_host="DH10B"
/note="Vector: PCMV-SPOrt6"
338.1840
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH16942.1"
/db_xref="GI:168773877"
/translation="MAYTKAPKKQIQPADDMOETKFPKTRSRISRSISOSSTDS
YSSASYSDDSDSEVSPREKQNTNSKSNFCVKNKQAEFRREIEIAODMSALIS
LRKRAOGKPLAGAKIVGCTHTAQAVALIETCALAGARMSACNYSYONEVAAL
LAAGVAVAMKESSEDDFMWICIDRCNMNDGMANMLIDGGDLTHWYKRYPNVEKI
RGIYESTYVGHRIQLSKAKLCPANMNVNBYKOKPNLYCRESIIDGLKRTD
VMFGKQYVCGYGVGEGKCCALAKALGAVITTEIDPICALACMDGFRVKNLEVT
ROVDVITCTGKRNVTREHLDNRKNSCIYCNMHSNTEIDVTSLEPRTWERNQ
VDHYIAPDGKRYVLLAEGRLNLSCSTVPFVLSIATTDQALLIELYNAPBGRYKOD
VYLPKRDDEVYASLHPFSDAHLTELTDQAKYLGINKNGPRPNRYR"
BASE COUNT 644 a 600 c 671 g 637 t
ORIGIN
Query Match 98.5%; Score 2525.2; DB 9; Length 2552;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2549; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
QY 8 GCAGGTGGAGCTCGAGCTGCTGCTTCTCTTGGCCACCGTCCCTGCCGC 67
DB 1 GCAGGTGGAGCTCGAGCTGCTGCTTCTCTTGGCCACCGTCCCTGCCGC 60
QY 68 TGCCTTGGCTGCGCAACAGACAGCGTGGCCACAGCACTCAGAAAGCGAGCAGCT 127
DB 61 TGCCTTGGCTGCGCAACAGACAGCGTGGCCACAGCACTCAGAAAGCGAGCAGCT 120
QY 128 CGAGCAGAGGGCCGCGAGAGGGTGGCGATCGCGTTCGAGAGGGCCGCGCGGAGG 187
DB 121 CGAGCAGAGGGCCGCGAGAGGGTGGCGATCGCGTTCGAGAGGGCCGCGCGGAGG 180
QY 188 CGGGGGGGGGCCAGAGGGGAAAGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 247
DB 181 CGGGGGGGGGCCAGAGGGGAAAGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 240
QY 248 CGGGGGAATGCGATGCTGACGCGATGCGCTGCGCGGGGTCGGGGAGAGCTGAACA 307
DB 241 CGGGGGAATGCGATGCTGACGCGATGCGCTGCGCGGGGTCGGGGAGAGCTGAACA 300
QY 308 GGCAGAGGAGTCGAGAGCGCCGAGAAATCTCTTCATGGCCACCGTCAACAGGCGCC 367
DB 301 GGCAGAGGAGTCGAGAGCGCCGAGAAATCTCTTCATGGCCACCGTCAACAGGCGCC 360
QY 368 CAAGAAGCAAAATCCATTTGCTGATGATGATGATGATGATGATGATGATGATGATG 427
DB 361 CAAGAAGCAAAATCCATTTGCTGATGATGATGATGATGATGATGATGATGATGATG 420
QY 428 TGGCCGAAGATCTTGTGCTGCTGATCTCAGAGTCTCCAGTACAGTACAGTACAGT 487
DB 421 TGGCCGAAGATCTTGTCTGCTGCTGATCTCAGAGTCTCCAGTACAGTACAGTACAGT 480
QY 488 TGCATCTTACACAGATAGCTGTATGATGATGATGATGATGATGATGATGATGATGAT 547
DB 481 TGCATCTTACACAGATAGCTGTATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 548 CTCGCAAGGGCAGAGCAATTTCTGTGAAGAAATCAAGAGGCAAAATTTGAGCGCG 607
DB 541 CTCGCAAGGGCAGAGCAATTTCTGTGAAGAAATCAAGAGGCAAAATTTGAGCGCG 600
QY 608 GAGAGTTGAGATTCAGAGCAAGATGCTGCTGATTTCACTCAGAGAAAGCTGTCA 667
DB 601 GAGAGTTGAGATTCAGAGCAAGATGCTGCTGATTTCACTCAGAGAAAGCTGTCA 660
QY 668 GGGGGAGAAAGCCCTGGCTGGCTAAATATGATGATGATGATGATGATGATGATGATG 727
DB 661 GGGGGAGAAAGCCCTGGCTGGCTAAATATGATGATGATGATGATGATGATGATGATG 720
QY 728 AGCGGTGTGATGAGACACTGTGTGCGCGGAGGAGGCGGCTGGCTGTGTA 787

DB 721 AGCGGTGTGATGAGACACTGTGTGCGCGGAGGAGGCGGCTGGCTGTGTA 780
QY 788 CATCTACTCACTCAGAGTGAAGTACGTGACAGACTGCTGAGGCTGAGTGTGAGTGT 847
DB 781 CATCTACTCACTCAGAGTGAAGTACGTGACAGACTGCTGAGGCTGAGTGTGAGTGT 840
QY 848 CCCTTGGAGGGCGAGTCAAGAAATGACTCTCTGCTGATGATGATGATGATGATGATGAT 907
DB 841 CCCTTGGAGGGCGAGTCAAGAAATGACTCTCTGCTGATGATGATGATGATGATGATGAT 900
QY 908 GAGTGGGTGGCAGGCGCAACATGATCCGATGATGATGATGATGATGATGATGATGATGAT 967
DB 901 GAGTGGGTGGCAGGCGCAACATGATCCGATGATGATGATGATGATGATGATGATGATGAT 960
QY 968 TAAGAATATCCAAAGCTGTTTAAAGATCCGAGGCTGTGGAAGAGAGCGTGAATG 1027
DB 961 TAAGAATATCCAAAGCTGTTTAAAGATCCGAGGCTGTGGAAGAGAGCGTGAATG 1020
QY 1028 TGTTCACAGGCTGTATCAGTCTTCCAAAGCTGGGAACTGTGTTCGGCCATGAAAGCT 1087
DB 1021 TGTTCACAGGCTGTATCAGTCTTCCAAAGCTGGGAACTGTGTTCGGCCATGAAAGCT 1080
QY 1088 CATGATCTGTTACCAACAGAAAGTTGATTAATCTGCTGCTGCGAGAAATCATTTT 1147
DB 1081 CATGATCTGTTACCAACAGAAAGTTGATTAATCTGCTGCTGCGAGAAATCATTTT 1140
QY 1148 GATGAGCTGGAAGAGACACAGATGATGATGATGATGATGATGATGATGATGATGATG 1207
DB 1141 GATGAGCTGGAAGAGACACAGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1208 CTATGAGTGAAGTGAAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267
DB 1201 CTATGAGTGAAGTGAAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1268 CATTAACCAAAATGACCCCATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1327
DB 1261 CATTAACCAAAATGACCCCATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1328 AAAGCTAAATGAATGATCCGCGCAAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1387
DB 1321 AAAGCTAAATGAATGATCCGCGCAAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
QY 1388 TGTAGTGAACGCGGAGAGTGTGATGCTGATGATGATGATGATGATGATGATGATGAT 1447
DB 1381 TGTAGTGAACGCGGAGAGTGTGATGCTGATGATGATGATGATGATGATGATGATGAT 1440
QY 1448 CCACTCCAACACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1507
DB 1441 CCACTCCAACACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
QY 1508 AGTACGTTCTCAGGTGACATGATGATGATGATGATGATGATGATGATGATGATGATG 1567
DB 1501 AGTACGTTCTCAGGTGACATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
QY 1568 AGAGGTCGCTACTCAATTTGAGCTGCTCCAGAGTCCCACTTGTGCTGCTGCTGCT 1627
DB 1561 AGAGGTCGCTACTCAATTTGAGCTGCTCCAGAGTCCCACTTGTGCTGCTGCTGCTGCT 1620
QY 1628 AGCCACAACACAGAGCTTGGCAGTGAATGATGATGATGATGATGATGATGATGATGATG 1687
DB 1621 AGCCACAACACAGAGCTTGGCAGTGAATGATGATGATGATGATGATGATGATGATGATG 1680
QY 1688 GCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1747
DB 1681 GCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
QY 1748 ATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1807
DB 1741 ATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
QY 1808 CAAAAATGGGCAATTCAAACCTATTTATTAAGATGATGATGATGATGATGATGATGATGAT 1867
DB 1801 CAAAAATGGGCAATTCAAACCTATTTATTAAGATGATGATGATGATGATGATGATGATGAT 1860

```

OY 1868 CAGTCACCTGAAACACACACTCTAAAGAAATATTTTAAAGATACTTTATTTCTTC 1927
DB 1861 CAGTCACCTGAAACACACACTCTAAAGAAATATTTTAAAGATACTTTATTTCTTC 1920
OY 1928 TTACACCTTCCCTGATTTTCTTCTATTAATTCATCTCTGTTTTTCTCTCATAT 1987
DB 1921 TTACACCTTCCCTGATTTTCTTCTATTAATTCATCTCTGTTTTTCTCTCATAT 1980
OY 1988 CCAAGTTCTGCAGACACACAGAGAACTGCTTCATGCTCTTTAGATGAATAGAGATTC 2047
DB 1981 CCAAGTTCTGCAGACACACAGAGAACTGCTTCATGCTCTTTAGATGAATAGAGATTC 2040
OY 2048 AGGTCCTCCTACTCTAGTCACTAAAGAGATTTTACTCCGCCAGCCAGAAAGGTAT 2107
DB 2041 AGGTCCTCCTACTCTAGTCACTAAAGAGATTTTACTCCGCCAGCCAGAAAGGTAT 2100
OY 2108 CTTCTCTTACATTTCTGGGACCTTAGTCTTAATAGGACTTATTAACAGGAATG 2167
DB 2101 CTT-TCCTTACATTTCTGGGACCTTAGTCTTAATAGGACTTATTAACAGGAATG 2159
OY 2168 CTAAGTACCTTCTCTGTGACATCTGCATGTCTAAATCGCCTTAAAGAGCCAT 2227
DB 2160 CTAAGTACCTTCTCTGTGACATCTGCATGTCTAAATCGCCTTAAAGAGCCAT 2219
OY 2228 TCTTACCTCTGAAATCAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2287
DB 2220 TCTTACCTCTGAAATCAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2279
OY 2288 CGGAGGTAGGTAGATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG 2347
DB 2280 CGGAGGTAGGTAGATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG 2339
OY 2348 TTTCCTGCTTAAAGTGTGTGTCTCAATCTGAGATGTGTATAGAGATGCTTGT 2407
DB 2340 TTTCCTGCTTAAAGTGTGTGTCTCAATCTGAGATGTGTATAGAGATGCTTGT 2399
OY 2408 AATCAGAGAGGCGGCACTTGGATTATATGATGATGATGATGATGATGATGATGATG 2467
DB 2400 AATCAGAGAGGCGGCACTTGGATTATATGATGATGATGATGATGATGATGATGATG 2459
OY 2468 TCAATTTTGAAGTTTAACTAGTCACTATTTAGATTTATTTTCTCTCTAGGA 2527
DB 2460 TCAATTTTGAAGTTTAACTAGTCACTATTTAGATTTATTTTCTCTCTAGGA 2518
OY 2528 TTTATTTCTGTGTCCCAAAAAA 2561
DB 2519 TTTATTTCTGTGTCCCAAAAAA 2552

RESULT 5
LOCUS BC007576 2526 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
ACCESSION BC007576
VERSION BC007576.1 GI:14043176
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 M (bases 1 to 2526)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk.
Email: cgapbs-rt@mail.nih.gov
Tissue Procurement: ATCC
REMARK COMMENT

```

```

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Gull,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natesja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Zetta.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 22 Row: h Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1. 2526
/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="locusid:10768"
/clone="MGC:15558 IMAGE:3139729"
/tissue_type="Placenta, choriochorionoma"
/clone_lib="NIH-MGC-21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
308..1810
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH07576.1"
/db_xref="gi:14043177"
/translation="MAVTYKPKKQLOFADMDQEFPTPTTGRSLRSISOSTDS
ISSASTSDSDVSPREKQJNSKSSNFCKANIKQAEGRLEIEAEDMSLIS
LRRAGGERFLAKGIVGCTHTTAQTAIVLLETICALAQCSACNISTONEAAL
AEGAVAFPMKGESEDFWMCIDRCVMDQAMMILDDGDLTHWYKPKPNFKKI
RGIRESYTVGHLIYOLSKAGKLCVPAMNYNDSTKQFIDLXCCRESILGLKRTD
VMEGKQVYVCGYEGVKGKCAALKALGALVYTEIDICALQCMGFRVYKNEVI
ROYDVITTTGKNNVYTRREHLDNRKNSCIYCMGHSSTEIDYSLRPELMEWRRSQ
VDHVIWPDGRVYVLDGRILNLSCTSPFVISTATQALALIELTNAEGRYKOD
VLLPKMDEYVASLHLPSPDAHLTETLDDQATYLGINKNGPFRPNYR"

BASE COUNT 644 a 592 c 660 g 630 t
ORIGIN

Query Match 97.4%; Score 2495.6; DB 9; Length 2526;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2520; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

OY 38 GTTCTTGGGCGACCGTGCCTGCTCCGGGCTTGAGGCTGCGCAACAGAGCGG 97
DB 1 GTTCTTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
OY 98 GGCACAGACCTCAGAGAGCGAGAGCTGAGCAGAGGCGCGGAGAGGCTGCGGA 157
DB 61 GGCACAGACCTCAGAGAGCGAGAGCTGAGCAGAGGCGCGGAGAGGCTGCGGA 120
OY 158 TCGCGTGTGGAGAGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 217
DB 121 TCGCGTGTGGAGAGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
OY 218 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 277
DB 181 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
OY 278 GGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 337
DB 241 GGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
OY 338 CTCCTTCATGCGCAGCGTACCAAGGCGCCCAAGAGCAATCCAGTTGCTGATGACAT 397

```

301 CTCCTTACATGGCCACCCGTCACCAAGGCGCCCAAGAACCAATCCAGTTTGGCTGATGACAT 360
398 GCAGAGATTACACAAATTCGCCACCAAACTGGCCGAAGATCTTTGTCGCTGATCTC 457
361 GCAGAGATTACACAAATTCGCCACCAAACTGGCCGAAGATCTTTGTCGCTGATCTC 420
458 AACAGCTCTCAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 517
421 AACAGCTCTCAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 480
518 GGTTCCTCCCGAG 577
481 GGTTCCTCCCGAG 540
578 GAACATCAAG 637
541 GAACATCAAG 600
638 TGCTCTGATTTCTACAG 697
601 TGCTCTGATTTCTACAG 660
698 AGAGGCTGTACACATCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 757
661 AGAGGCTGTACACATCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 720
758 GGGGGCTCAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 817
721 GGGGGCTCAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
818 AGACAGGCTGAG 877
781 AGACAGGCTGAG 840
878 CTGGGTGATTAAGACAG 937
841 CTGGGTGATTAAGACAG 900
938 TGATGGGGGAGACTTAACCCACTGGGTTTATAGAGATATCCAAACGTGTTTAAAGAGAT 997
901 TGATGGGGGAGACTTAACCCACTGGGTTTATAGAGATATCCAAACGTGTTTAAAGAGAT 960
998 CCGAGGCTATGTTGAG 1057
961 CCGAGGCTATGTTGAG 1020
1058 TGGGAGAGCTGTGTTCCGCGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1117
1021 TGGGAGAGCTGTGTTCCGCGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1118 TAACCTGATCTGCTGCCAGAGATCCATTTTGGATGGCTGAGAGAGAGAGAGAGAGATGAT 1177
1081 TAACCTGATCTGCTGCCAGAGATCCATTTTGGATGGCTGAGAGAGAGAGAGAGATGAT 1140
1178 GTTGGTGGGAAACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1237
1141 GTTGGTGGGAAACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1200
1238 TGCTCAAGAGCTGTGGAGAGATGCTATACATACGAGAGAGAGAGAGAGAGAGAGAGAG 1297
1201 TGCTCAAGAGCTGTGGAGAGATGCTATACATACGAGAGAGAGAGAGAGAGAGAGAGAG 1260
1298 GCAGGCTGATGAG 1357
1261 GCAGGCTGATGAG 1320
1358 TGTCGTAATTAACCTGACAG 1417
1321 TGTCGTAATTAACCTGACAG 1380
1418 GAAAGAGAGTGTATGCTATGAG 1477
1381 GAAAGAGAGTGTATGCTATGAG 1440

1478 CTCCTGACACTCCGAG 1537
1441 CTCCTGACACTCCGAG 1500
1538 GCCAGATGGGAAAG 1597
1501 GCCAGATGGGAAAG 1560
1598 CACAGTTCCACCTTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1657
1561 CACAGTTCCACCTTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
1658 ACTCTAATTAAGACACCCGAG 1717
1621 ACTCTAATTAAGACACCCGAG 1680
1718 GGATGAATACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1777
1681 GGATGAATACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
1778 AGATGACCAACCAAAATATCTGGAGCTCAACCAAAATGGGCAATTCACCTAATTAATTA 1837
1741 AGATGACCAACCAAAATATCTGGAGCTCAACCAAAATGGGCAATTCACCTAATTAATTA 1800
1838 CAGATTAATTAAGAC 1897
1801 CAGATTAATTAAGAC 1860
1898 ATATTTTAAAGATTAACCTTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1957
1861 ATATTTTAAAGATTAACCTTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1920
1958 AATTTCAATCTTTGTTTTCATCTTATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 2017
1921 AATTTCAATCTTTGTTTTCATCTTATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1980
2018 TTCAATGCTCTTTAGATTAATTAAGAGTTCAAGGTCCTCCTCCTCCTCCTCCTCCTCCTCCT 2077
1981 TTCAATGCTCTTTAGATTAATTAAGAGTTCAAGGTCCTCCTCCTCCTCCTCCTCCTCCTCCT 2040
2078 ATTTTACCTCCCGAG 2137
2041 ATTTTACCTCCCGAG 2099
2138 CTTAATTAAGTACCTTATTAACAGAGAAATGCTAAGGTACCTTCTGCTGAGAACATCTGC 2197
2100 CTTAATTAAGTACCTTATTAACAGAGAAATGCTAAGGTACCTTCTGCTGAGAACATCTGC 2159
2198 AATGCTAATTCGCTTAAAG 2257
2160 AATGCTAATTCGCTTAAAG 2219
2258 TTCTTCAG 2317
2220 TTCTTCAG 2279
2318 GTTAATTTCCCTTAAAG 2377
2280 GTTAATTTCCCTTAAAG 2339
2378 AGAGATGTAATTAAG 2437
2340 AGAGATGTAATTAAG 2399
2438 TAGCCCTTCGCTCCACCTCCACAG 2497
2400 TAGCCCTTCGCTCCACCTCCACAG 2459
2498 TCTAATTTAGTTTAAATTTGCTCTAGAGATTAATTTGCTCTAGAGATTAATTTGCTCTAGAGAT 2557
2460 TCTA-TTGAATTAATTTGCTCTAGAGATTAATTTGCTCTAGAGATTAATTTGCTCTAGAGAT 2518

QY 2558 AAAAA 2563
 DB 2519 AAAAA 2524

RESULT 6
 HS082761. 2258 bp mRNA linear PRI 17-DEC-2001
 LOCUS HS082761. Homo sapiens s-adenosyl homocysteine hydrolase homolog (XPVKona)
 DEFINITION
 mRNA, complete cds.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 1 (bases 1 to 2258)
 Cleaver, J.E., Afzal, V., Feeney, L., McDowell, M., Sadlinski, W.,
 Volpe, J.P.G., Busch, D.B., Coleman, D.M., Ziffer, D.W., Yu, Y.,
 Nagasawa, H. and Little, J.B.
 Increased ultraviolet sensitivity and chromosomal instability
 related to p53 function in the xeroderma pigmentosum variant
 Cancer Res. 59 (5), 1102-1108 (1999)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 99188317
 2 (bases 1 to 2258)
 Volpe, J.P.G., McDowell, M., Jostes, R.F., Afzal, V., Sadlinski, W.,
 Trask, B.J., Legerski, R. and Cleaver, J.E.
 Complementation of chromosomal instability in the xeroderma
 pigmentosum variant by a gene on human chromosome 1 with homology
 to S-adenosyl homocysteine hydrolase
 Unpublished
 3 (bases 1 to 2258)
 Volpe, J.P.G., McDowell, M. and Cleaver, J.E.
 Direct Substitution
 Submitted (19-DEC-1996) Dermatology, UCSF, 3rd and Parnassus, Box
 0750, San Francisco, CA 94143, USA
 Location/Qualifiers
 1. 2258
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /cell_type="fibroblasts"
 /tissue_type="skin"
 1. 2258
 /gene="XPVKona"
 48. 1550
 /note="xeroderma pigmentosum variant"
 /product="S-adenosyl homocysteine hydrolase homolog"
 /protein_id="AAC01960.1"
 /db_xref="GI:2852125"
 /translation="MATVTKAPKKQIOFADDMQEFTEPKTKGRSLRSISQSTDS
 YSSAAYSDSDDEVSPEKQNTSKSGSNFCVKNIOAQEFGREIEIAEQDMALIS
 LKRAQGEKPLAGAKVCTHITAQTAVALLETICALAQGMSACNISTQENYAAAL
 AEAGAVAFAMKGESEDDFMWCIDRCVNMGMOMAMILDGDDLHWYKXKPNPKKI
 RGIYESVTGVRITOLSKAGIKCPAMVNDSTVKQKPNKXCCRESIIDGLKRTD
 VAFGKQYVCGTGEYKGCCKCALKAIGALVITTEIDPICALACMGFRVYKLNKVI
 ROVDVITCTGNKNVTRREHLDMKNKSCIVCNMGSHTETIDVSLRPELTWERVRSQ
 VDHTVTPDKRYVLAEGRLNLSCTVPEFVLSITATTOALALIELYNAEGRYKOD
 VYLPERKMEYVASLHLPSFDAHLTELTDQAKYLGNKNGPFKNRYR"

BASE COUNT 603 a 513 c 542 g 600 t
 ORIGIN

Query Match 86.6%; Score 2218.6; DB 9; Length 2258;
 Best Local Similarity 99.7%; P-Value: 0;
 Matches 2254; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY 298-ACCTGAAGAGCGCAAGAGATGAGAGAGCCGAGAGTACTCTTCTATGGCCACCGCA 357
 DB 1-ACCTGAAGAGCGCAAGAGATGAGAGAGCCGAGAGTACTCTTCTATGGCCACCGCA 60

QY 358 CCAAGGCGCCCAAGAGCAAAATCCAGTTTGTGATGACATGAGAGAGTCCACCAATTC 417
 DB 61 CCAAGGCGCCCAAGAGCAAAATCCAGTTTGTGATGACATGAGAGAGTCCACCAATTC 120
 QY 418 CCACCAAACTGGCCGAAGATCTTGTCTCGCTGCATCTCAAGTCTCCACATGACAT 477
 DB 121 CCACCAAACTGGCCGAAGATCTTGTCTCGCTGCATCTCAAGTCTCCACATGACAT 180
 QY 478 ACAGTTAGCTGCATCTCCACAGATAGCTGATGATGATGATGATGATGATGATGATG 537
 DB 181 ACAGTTAGCTGCATCTCCACAGATAGCTGATGATGATGATGATGATGATGATGATG 240
 QY 538 AGCAAACTCCCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597
 DB 241 AGCAAACTCCCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 QY 598 TTGAGAGCGCGGAGATGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 657
 DB 301 TTGAGAGCGCGGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 QY 658 AACGTGCTCAGGGGAT 717
 DB 361 AACGTGCTCAGGGGAT 420
 QY 718 CAGCCAG 777
 DB 421 CAGCCAG 480
 QY 778 CTGCTTTGATCTTCTCAACTCAAGATGAGATGAGATGAGATGAGATGAGATGAG 837
 DB 481 CTGCTTTGATCTTCTCAACTCAAGATGAGATGAGATGAGATGAGATGAGATGAG 540
 QY 838 TTGAGAGTTCGCTTGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897
 DB 541 TTGAGAGTTCGCTTGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 QY 898 GTGGAACATGATGAG 957
 DB 601 GTGGAACATGATGAG 660
 QY 958 ACTGGGTTTATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1017
 DB 661 ACTGGGTTTATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
 QY 1018 GCGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1077
 DB 721 GCGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
 QY 1078 CCATGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1137
 DB 781 CCATGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
 QY 1138 AATCCATTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1197
 DB 841 AATCCATTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 QY 1198 TGGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1257
 DB 901 TGGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
 QY 1258 CAATTTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1317
 DB 961 CAATTTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 QY 1318 TCAGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1377
 DB 1021 TCAGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 QY 1378 GAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1437
 DB 1081 GAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140


```

OY 1438 GCATATGGGGCCACTCCACAGAAATGATGTGACCCAGCTCCGACCTCCGGAGCTGA 1497
DB 1141 GCAATATGGGGCCACTCCACAGAAATGATGTGACCCAGCTCCGACCTCCGGAGCTGA 1200
OY 1498 CGTGGAGAGGAGTACGTTCTCAGGTGAGACATGTCTGGCCAGATGGCAAGAGTTG 1557
DB 1201 CGTGGAGAGGAGTACGTTCTCAGGTGAGACATGTCTGGCCAGATGGCAAGAGTTG 1260
OY 1558 TCCTCTGGCAGAGGGGTCTCTACTCAATTTGAGTGTCTCCACAGTTTCCACCTTTGTTTC 1617
DB 1261 TCCTCTGGCAGAGGGGTCTCTACTCAATTTGAGTGTCTCCACAGTTTCCACCTTTGTTTC 1320
OY 1618 TGTCCATCAGACCCACCAACAGAGGCTTTGGCAGTGTGATGAACTGTATTAATCCACCCGAGG 1677
DB 1321 TGTCCATCAGACCCACCAACAGAGGCTTTGGCAGTGTGATGAACTGTATTAATCCACCCGAGG 1380
OY 1678 GGCAGTACAGAGAGAGTGTACTGTCTTCTTAAGAAAATGATGATAGCTTCCAGACT 1737
DB 1381 GGCAGTACAGAGAGAGTGTACTGTCTTCTTAAGAAAATGATGATAGCTTCCAGACT 1440
OY 1738 TGCATCTGCCATCATTTGATGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATC 1797
DB 1441 TGCATCTGCCATCATTTGATGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATC 1500
OY 1798 TGGGACTCAACAAAATGGGCGCATCAAACTATTTACAGATACCTAATGGACCATAC 1857
DB 1501 TGGGACTCAACAAAATGGGCGCATCAAACTATTTACAGATACCTAATGGACCATAC 1560
OY 1858 TACCAAGAGACAGATCCACCTGAAACACACACACTCTAAGAAAATATTTTAAATTAATTT 1917
DB 1561 TACCAAGAGACAGATCCACCTGAAACACACACACTCTAAGAAAATATTTTAAATTAATTT 1620
OY 1918 TATTTTCTCTACTACCTCTTCTCTGTGATTTTCTCTATATTTCAATCTGTGTTTTC 1977
DB 1621 TATTTTCTCTACTACCTCTTCTCTGTGATTTTCTCTATATTTCAATCTGTGTTTTC 1680
OY 1978 ACTCATTTTCAAGTTCGACAGACACAGAGACTGTCTTCAAGTCTTGTAGATGA 2037
DB 1681 ACTCATTTTCAAGTTCGACAGACACAGAGACTGTCTTCAAGTCTTGTAGATGA 1740
OY 2038 ATAGAAGTTCAGAGGTCCTCTACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCAG 2097
DB 1741 ATAGAAGTTCAGAGGTCCTCTACTCTAGTCACTAAGAAAGATTTTACTCCCCAGCCAG 1800
OY 2098 AAGGTGATTTCTCTCTTACATTTCTGGGAGTATGTTTATTAATTAATTAATTAATTA 2157
DB 1801 AAGGTGATTTCTT-TCTTACATTTCTGGGAGTATGTTTATTAATTAATTAATTAATTA 1859
OY 2158 AAGGAAATGCTAAGTACTCTCTGTGGAACAATCTGCAATGCTAATAATCGCCTTAA 2217
DB 1860 AAGGAAATGCTAAGTACTCTCTGTGGAACAATCTGCAATGCTAATAATCGCCTTAA 1919
OY 2218 AAGGAAATGCTAAGTACTCTCTGTGGAACAATCTGCAATGCTAATAATCGCCTTAA 2277
DB 1920 AAGGAAATGCTAAGTACTCTCTGTGGAACAATCTGCAATGCTAATAATCGCCTTAA 1979
OY 2278 GGTACTACCCGAGAGTGTAGATGTGGTGTGATGTGATTTTCTCTTGAAGT 2337
DB 1980 GGTACTACCCGAGAGTGTAGATGTGGTGTGATGTGATTTTCTCTTGAAGT 2039
OY 2338 CCAAGCCCTGTTTCTGCGTAAAGTGTGATGTGCAATGCTGAGAGATGTATTAATGACA 2397
DB 2040 CCAAGCCCTGTTTCTGCGTAAAGTGTGATGTGCAATGCTGAGAGATGTG-ATAATGACA 2098
OY 2398 TGGCTTTGTAAGTACAGAGGCGCCACTGATTTATAGTAAAGCCCTCCCACTCCCA 2457
DB 2099 TGGCTTTGTAAGTACAGAGGCGCCACTGATTTATAGTAAAGCCCTCCCACTCCCA 2158
OY 2458 CCAGACTTCTCTCTTTTGAAGTTTAACTAGACACACTCTAATTTGAGTTTAAATTTG 2517
DB 2159 CCAGACTTCTCTCTTTTGAAGTTTAACTAGACACACTCTAATTTGAGTTTAAATTTG 2217
OY 2518 TCCTCTAGAGATTTATTTCTGTGTCCAAAAA 2558

```

DB 2218 TCCTCTAGAGATTTATTTCTGTGTCCAAAAA 2258

RESULT 7

LOCUS HSM800298

DEFINITION Homo sapiens mRNA; cDNA DKFZ564A1523 (from clone DKFZ564A1523);

ACCESSION AL049954

VERSION AL049954.1

KEYWORDS GI:4884203

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS 1 (bases 1 to 2510).

TITLE Wambut, R., Heuber, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

JOURNAL Direct Submission

COMMENT Submitted (15-MAY-1999) MIPS, Am Kiofierspitz 18a, D-82152

Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;

sequenced by AGO (Berlin/Germany) within the cDNA sequencing

consortium of the German Genome Project.

This clone (DKFZ564A1523) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heuberweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de; Further

information about the clone and the sequencing project is available

at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

Location/Qualifiers

1..2510

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="1, 15.55 cR from DIS248"

/clone="DKFZ564A1523"

/tissue_type="brain"

/clone_1kb="564 (synonym: hfbz2). Vector pAMP1; host

XI-2blue; sites NotI + SalI"

/dev_stage="fetal"

1..1796

/gene="DKFZ564A1523"

<1..1796

/note="S-adenosyl homocysteine hydrolase homolog XPYKona

splice variant"

/codon_start=3

/product="hypothetical protein"

/protein_id="CAB4323.1"

polyA site 664 a 556 c 618 g 672 t

BASE COUNT 664 a 556 c 618 g 672 t

ORIGIN

Query Match 84.3%; Score 2161.2; DB 9; Length 2510;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2185; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

OY 374 GCAATCCAGTTTGCATGACATGACAGGATCCACCAATTCGCCCAAACTGGCG 433

DB 323 GCAATCCAGTTTGCATGACATGACAGGATCCACCAATTCGCCCAAACTGGCG 382

434 AGATCTTGTGCTGATCTCAGTCTCCACTGACAGTCAAGTTCAGTCTCATC 493
|||||
383 AAGATCTTGTGCTGATCTCAGTCTCCACTGACAGTCAAGTTCAGTCTCATC 442
|||||
494 CTACAGATAGTCTGATGATGAGTTTCCCGAGAGAACAGCAACCACTCCAA 553
|||||
443 CTACAGATAGTCTGATGATGAGTTTCCCGAGAGAACAGCAACCACTCCAA 502
|||||
554 GGGCAGCAGCAATTTCTGTGAGAGACATCAAGCAGCAATTTGAGCCCGGAGAT 613
|||||
503 GGGCAGCAGCAATTTCTGTGAGAGACATCAAGCAGCAATTTGAGCCCGGAGAT 562
|||||
614 TGAGATGACAGAGAGATGTCGTCTGATTTCACTCAGAGAAACGCTCAGGGGGA 673
|||||
563 TGAGATGACAGAGAGATGTCGTCTGATTTCACTCAGAGAAACGCTCAGGGGGA 622
|||||
674 GAACCCCTTGGCTGCTAAATAATAGTGGGCTGTACACACATCACAGCCAGACGGT 733
|||||
623 GAACCCCTTGGCTGCTAAATAATAGTGGGCTGTACACACATCACAGCCAGACGGT 682
|||||
734 GTTATTGAGACACTCTGTGCTGCTGGGGGCTAGTCCGCTGCTGCTGTAACTGA 793
|||||
683 GTTATTGAGACACTCTGTGCTGCTGGGGGCTAGTCCGCTGCTGCTGTAACTGA 742
|||||
794 CTCAACTCAGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 853
|||||
743 CTCAACTCAGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 802
|||||
854 GAAGGGGAGTCAAGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 913
|||||
803 GAAGGGGAGTCAAGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 862
|||||
914 GTGGCAGGCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 973
|||||
863 GTGGCAGGCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 922
|||||
974 GTATCCAAACGCTGTTAAGAGATCCGAGCATTTGTGAGAGAGAGGCTGATGATG 1033
|||||
923 GTATCCAAACGCTGTTAAGAGATCCGAGCATTTGTGAGAGAGAGGCTGATGATG 982
|||||
1034 CAGGCTGATCAGCTCTCCAAAGCTGGGAGCTGTGTGCTGCTGCTGCTGCTGCTG 1093
|||||
983 CAGGCTGATCAGCTCTCCAAAGCTGGGAGCTGTGTGCTGCTGCTGCTGCTGCTG 1042
|||||
1094 TTTCTTACCAAAAGAGATTTGATTAATTTGATCTGCTGCTGCTGCTGCTGCTGCTG 1153
|||||
1043 TTTCTTACCAAAAGAGATTTGATTAATTTGATCTGCTGCTGCTGCTGCTGCTGCTG 1102
|||||
1154 CCTGAAGAGACACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1213
|||||
1103 CCTGAAGAGACACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1162
|||||
1214 TGAGGTAGGCAAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1273
|||||
1163 TGAGGTAGGCAAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1222
|||||
1274 CGAAATGACCCCATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1333
|||||
1223 CGAAATGACCCCATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1282
|||||
1334 AATGATGATCGGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1393
|||||
1283 AATGATGATCGGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1342
|||||
1343 GACACGGGAGCACTTGATTCGATGAAAAAGTGTGATGATGATGATGATGATGATG 1402
|||||
1394 GACACGGGAGCACTTGATTCGATGAAAAAGTGTGATGATGATGATGATGATGATG 1453
|||||
1454 CACACAGAAATCATGATGACACAGCTCCGCACTCCGCACTCCGCACTCCGCACTCCG 1513
|||||
1403 CACACAGAAATCATGATGACACAGCTCCGCACTCCGCACTCCGCACTCCGCACTCCG 1462
|||||
1514 TTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1573
|||||

Db 1463 TTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1522
|||||
Qy 1574 TCGTCTACCAATTTGAGCTGCTCCACAGTTCACACCTTTGCTGCTCATCAGCCAC 1633
|||||
Db 1523 TCGTCTACCAATTTGAGCTGCTCCACAGTTCACACCTTTGCTGCTCATCAGCCAC 1582
|||||
Qy 1634 AACACAGCTTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1693
|||||
Db 1583 AACACAGCTTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1642
|||||
Qy 1694 TGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1753
|||||
Db 1643 TGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1702
|||||
Qy 1754 TGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATTTCTGGACCTCAACAAA 1813
|||||
Db 1703 TGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATTTCTGGACCTCAACAAA 1762
|||||
Qy 1814 TGGGCCATTCACACCTAATTTTACAGTACTAATGACACTACTACCAAGACAGTCC 1873
|||||
Db 1763 TGGGCCATTCACACCTAATTTTACAGTACTAATGACACTACTACCAAGACAGTCC 1822
|||||
Qy 1874 ACCTGAACACACACTCTAAAGAAATTTTAAAGATTAATTTTAACTTTTCTCTACTC 1933
|||||
Db 1823 ACCTGAACACACACTCTAAAGAAATTTTAAAGATTAATTTTAACTTTTCTCTACTC 1882
|||||
Qy 1934 CTTTCTCTGATTTTCTCTAATTTTCAATTTCTGTTTCTGATTTCTGATTTCTGAT 1993
|||||
Db 1883 CTTTCTCTGATTTTCTCTAATTTTCAATTTCTGTTTCTGATTTCTGATTTCTGAT 1942
|||||
Qy 1994 TCTGCACACACACAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2053
|||||
Db 1943 TCTGCACACACACAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2002
|||||
Qy 2054 CTTCTACTGATCTCTAAAGAGATTTTACTCCCTCCAGCCAGAGAGGATGATTTCTCTC 2113
|||||
Db 2003 CTTCTACTGATCTCTAAAGAGATTTTACTCCCTCCAGCCAGAGAGGATGATTTCTT -TC 2061
|||||
Qy 2114 TTTACCAATTTCTGGGACTTTAGTCTTAATTTAGTACTTATTAACAGAGAAATGCTAACG 2173
|||||
Db 2062 TTTACCAATTTCTGGGACTTTAGTCTTAATTTAGTACTTATTAACAGAGAAATGCTAACG 2121
|||||
Qy 2174 TACCTTCTGCTGGAACAATGTCGAATGTCGAATGTCGAATGTCGAATGTCGAATG 2233
|||||
Db 2122 TACCTTCTGCTGGAACAATGTCGAATGTCGAATGTCGAATGTCGAATGTCGAATG 2181
|||||
Qy 2234 CTGCTGAATCAGTCTCTTCACTTCTTCAAGAGAGAGGATGATCTACCCGGCAG 2293
|||||
Db 2182 CTGCTGAATCAGTCTCTTCACTTCTTCAAGAGAGAGGATGATCTACCCGGCAG 2241
|||||
Qy 2294 GTAAGTTAGATGTTGGTGTGATGTTAATTTCCCTTGAAGTTCCAAAGCCCTGTTCT 2353
|||||
Db 2242 GTAAGTTAGATGTTGGTGTGATGTTAATTTCCCTTGAAGTTCCAAAGCCCTGTTCT 2301
|||||
Qy 2354 GCGTAAAGGTGATGTCACAGTTCAGAGATGTCATGATGATGATGATGATGATGATG 2413
|||||
Db 2302 GCGTAAAGGTGATGTCACAGTTCAGAGATGTCATGATGATGATGATGATGATGATG 2361
|||||
Qy 2414 GGAAGGCCACCTGATTTAGTATGATGATGATGATGATGATGATGATGATGATGATG 2473
|||||
Db 2362 GGAAGGCCACCTGATTTAGTATGATGATGATGATGATGATGATGATGATGATGATG 2421
|||||
Qy 2474 TTTGAGTTTAACTAGACTCACTTAATTTGAGTTTAAATTTTGTCTGATGATTTAT 2533
|||||
Db 2422 TTTGAGTTTAACTAGACTCACTTAATTTGAGTTTAAATTTTGTCTGATGATTTAT 2480
|||||
Qy 2534 TCTGTTTCCAAAAAATTTTAAAAA 2563
|||||
Db 2481 TCTGTTTCCAAAAAATTTTAAAAA 2510
|||||

RESULT 8

	LOCUS	BC018218	1772 bp	mRNA	linear	ROD 07-AUG-2002
	DEFINITION	Mus musculus, S-adenosylhomocysteine hydrolase-like 1, clone MGC:18748 IMAGE:4007102, mRNA, complete cds.				
	ACCESSION	BC018218				
	VERSION	BC018218.1	GI:17390492			
	KEYWORDS	MGC.				
	SOURCE	house mouse.				
	ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1772)				
	REFERENCE	Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
	AUTHORS	Stausberg, R.				
	JOURNAL	Direct Submission				
	REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
	COMMENT	Contact: MGC help desk Email: gcgaps@email.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amgbcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M., Yoon, V.S., Koyls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.				
	FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAX Plate: 23 Row: n Column: 6 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency OR analysis, Genomescan gene prediction, Similarity but not identity to protein.				
	SOURCE	location/Qualifiers 1. 1772 /organism="Mus musculus" /db_xref="taxon:10090" /map="CZECH II" /clone="MGC:18748 IMAGE:4007102" /tissue_type="Mammary tumor metastasized to lung." MMTV-LTR/nrl model. Expression driven by an MMTV-LTR enhancer. /clone_1lb="NCI CGAP Lu30" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 191. 1693 /codon_start=1 /product="S-adenosylhomocysteine hydrolase-like 1" /protein_id="AAH18218.1" /db_xref="GI:17390493" /db_xref="locusid:10768" /translation="MATWTKAPKKOIRADPOADMEFTKPTKTGRSLRSISQSSDPS YSSASITSDSDDEVSPEPKOTSKSGSNVCYNINIOAEGRREIELAEQDSALIS LRKRAQGEPLAGAKIVGCTHTAQAVLITFLALGAQCWSACNTISTONEVTAALAI AEADVAYFAVMRGESDEDSPKICLVPAENVNDSTVKOFDMLYCCRESILDLGRTHDI RGIIESEVTGAMRLEYOLSKFKLCVPANVNDSTVKOFDMLYCCRESILDLGRTHDI VMFGKOYVVGEGEVGGCAALKALCALAIYITEIDPICLAQMDFRYVAKLENAVYIA RDVDVITCGNKNVYIREHLDPMKNKSIVCNMGMSNEIDVTSLSRPELWEVERSCS VDHYTPDPGRKVYVALBRGLNISCSTPYFVLSTLTATOTALILEYNAEGRRKOD VYLLEKMDDEVYSLHLPDSFAHLTEIDDQAKYIGLNKNPFRPNRYTY"				
	CDS	BASE COUNT 462 a 406 c 506 g 398 t				
	ORIGIN	Query Match 59.5%; Score 1525.4; DB 10; Length 1772; Best Local Similarity 92.7%; Pred No. 7e-299; Matches 1635; Conservative 0; Mismatches 121; Indels 7; Gaps 3				

OY	152	GGGCGATGGCGGTGTGGAGAGGGCCGCGCGGGCAGGGCGGGCGCCGACGAGAGGGGAAG	211
Db	2	GGGCGAGCGCGGTGGCCGAGAGCGCCGCGCGG----	57
OY	212	AGGCGGGGGCGGCGGGATGAGCGCGCTGGCGCGGGCCGGGGGGAATGTGATGCTGACGC	271
OY	272	GATGCCCTGCCCCGGGGGTGCGGGGAGGAGCTGAAGCAAGGCCAAGAGATGAGAGACCCGA	331
Db	118	GATGCCGCGTCCCGGTGTGCGGGGAGAGGTGAACAGGGCCAAAGGAATGAGAGACCCGA	177
OY	332	GAAGTACTCTTCATAGGCGCACCGTCACACAAAGCGCCCAAGAAACAATCAGTTTGCTGA	391
Db	178	GAAGTACTCTTCATAGGCGCACCGTCAACCAAGGCTCCCAAGAAACAATCAGTTTGCTGA	237
OY	392	TGACATGCAAGAGTTCACCAAAATTCGCCACCAAACTGGCCGAGATCTTTGTCTGCTC	451
Db	238	TGACATGCAAGAGTTCACCAAAATTCCTCTCTAAGACTGGCCGAGATCTTTGTCTGCTC	297
OY	452	GATCTCACAGTCTCTCACAGTACAGAGTTCAGTGCATGCTTACACAGATAGCTCTGA	511
Db	298	CATCTCACATCTCTCACAGACAGCTACAGTTCAGTGCATCTTATACAGATAGCTCTGA	357
OY	512	TGATGAGTCTTCCCGCGAGAGAGCAGCAAAACAATCCCAAGGCGACGCAATTTCTG	571
Db	358	TGATGAGTCTTCCCGCGAGAGAGCAGCAAAACAATCCCAAGGCGACGCAATTTCTG	417
OY	572	TGTGAAGAACATCAGACAGGCGAGAAATTTGAGCGCCGGGAGATTGAGATTTCAGACAGA	631
Db	418	TGTGAAGAACATCAGACAGGCGAGAAATTTGAGCGCCGGGAGATTGAGATTTCAGACAGA	477
OY	632	CATGCTGCTGTATTTCACTCAGAAACGCTGTCAGGGGGAGAAAGCCCTTGCTGGCTGC	691
Db	478	CATGCTGCTGTATTTCACTCAGAAACGCTGTCAGGGGGAGAAAGCCCTTGCTGGCTGC	537
OY	692	TAAATATGTGGGCTGTACACATCAACGCCAGACAGCGGTGTGATTGAGACACTCTG	751
Db	538	TAAATATGTGGGCTGTACACATCAACGCCCGACAGCGGTGTGATTGAGACACTCTG	597
OY	752	TGCGCGGGGGCGCATGCGCGCGGTGCTGTTTAACATCTACTCACTCAGATGAGT	811
Db	598	TGCGCGGGGGCGCATGCGCGGTGCTGCTGCACATCTATTCACTCAGATGAGT	657
OY	812	AGCTGAGCAGCTGGCTGAGGCGTGGAGTTGCAGTTCCTGGAAGGGCGAGTCAAGA	871
Db	658	AGCTGAGCAGCTGGCTGAGGCGTGGAGTTCCTGGAAGGGCGAGTCAAGA	717
OY	872	TGACTTCTGCTGTATTGATTCACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGAT	931
Db	718	TGACTTCTGCTGTGTATTGATTCACCGCTGTGTGAACATGATGGGTGGCAGGCCAATGAT	777
OY	932	CCGAGATGATGGGGAGAGACTTAACCACTGGTTTATAAAGATATCAACAGTGTAA	991
Db	778	CCGAGATGATGGGGAGAGACTTAACCACTGGTTTATAAAGATATCAACAGTGTAA	837
OY	992	GAGATCCGAGGCAATTTGGAGAGAGAGCGTGAAGCTGTTCAAGCGCTGTATCAGCTCTC	1051
Db	838	GAGATCCGAGGCAATTTGGAGAGAGAGCGTGAAGCTGTTCAAGCGCTGTATCAGCTCTC	897
OY	1052	CAAGCTGGGAACCTGTGTTCCGGCAGGAACGCTAAAGATTCCTTCCAAACGAA	1111
Db	898	CAAGCTGGGAACCTGTGTTCCAGCAGGAATGTAAGATTCCTTACCAACGAA	957
OY	1112	GTTTGAATACCTTACTGCTGCCGAGATCCATTTTGGATGGCCTGAAGAGACACAGA	1171
Db	958	GTTTGAATACCTTACTGCTGCCGAGATCCATTTTGGATGGCCTGAAGAGACACAGA	1017
OY	1172	TGTGATGTTTGGTGGGAAACAGTGGTGGTGTGGCTATGTGATAGGTATAGAGGGCTG	1231
Db	1018	TGTGATGTTTGGTGGGAAACAGTGGTGGTGTGGCTATGTGATAGGTATAGAGAGGGCTG	1077

1801 GACTCAACAAAAATGGGGCATTCACAACTAATATATACAGTACTAATGACCAATACAC 1860
 611 GACTCAACAAAAATGGGGCATTCACAACTAATATATACAGTACTAATGACCAATACAC 670
 1861 CAAGGACCACTGACCTGAAACACACACTCTAAAGAAATATTTTAAAGAACTTTAT 1920
 671 CAAGGACCACTGACCTGAAACACACACTCTAAAGAAATATTTTAAAGAACTTTAT 730
 1921 TTTCTCTTCTACCT 1980
 731 TTTCTCTTCTACCT 790
 1981 TCATTTTCCAAAGTTCTGACAGACACACAGAACTTCTCTCTCTCTCTCTCTCTCT 2040
 791 TCATTTTCCAAAGTTCTGACAGACACACAGAACTTCTCTCTCTCTCTCTCTCTCT 850
 2041 GAAGTTCAGGGGCTCTACCT 2100
 851 GAAGTTCAGGGGCTCTACCT 910
 2101 GGTGATTTCT 2160
 911 GGTGATTTCT 969
 2161 GGAATGCTTACCT 2220
 970 GGAATGCTTACCT 1029
 2221 GCCATTTCT 2280
 1030 GCCATTTCT 1089
 2281 ACCTACCCGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
 1090 ACCTACCCGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1149
 2341 AGCCCTGTTCT 2400
 1150 AGCCCTGTTCT 1209
 2401 CTTGTAAGATGAGAGGCGCCACTGTGATGATGATGATGATGATGATGATGATGATG 2460
 1210 CTTGTAAGATGAGAGGCGCCACTGTGATGATGATGATGATGATGATGATGATGATG 1269
 2461 GACTTCT 2520
 1270 GACTTCT 1328
 2521 TCTAGATTTATTTCTGTTGTCACAAAAA 2561
 1329 TCTAGATTTATTTCTGTTGTCACAAAAA 1369

RESULT 10
 BC003631 1323 bp mRNA linear PRI 12-JUL-2001
 LOCUS BC003631 1323 bp mRNA linear PRI 12-JUL-2001
 DEFINITION Home sapiens, similar to s-adenosylhomocysteine hydrolase-like 1,
 clone IMAGE:3010755, mRNA, partial cds.
 ACCESSION BC003631
 VERSION BC003631.1 GI:13177702
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1323)
 Strausberg, R.
 Direct Submission
 Submitted (26-FEB-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadan@systemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettelman and Anuradha Madan

COMMENT

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 3 Row: D Column: 6.
 Location/Qualifiers

FEATURES

source

CDS

1..1323
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3010755"
 /tissue_type="Muscle, thymomyosarcoma"
 /clone_id="NIH_MGC_17"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 <1..611
 /codon_start=3
 /product="Similar to s-adenosylhomocysteine hydrolase-like
 1"
 /protein_id="AAH03631.1"
 /db_xref="GI:13177703"
 /translation="HEBALGAIYITTEIDPICALQACMGFRVYKLEIVRODVVIT
 GKNKNVATREHIDRMKNSCTVCMGHSNTEIDVTSIRKTELWERKRSVDHYIWD
 GRVYLLAGRLNLSCSTVPFVLSITATQALALIEIENAEGRKDDVYLLPKRM
 DEYVASIHPSPDALHTEIYTDQAKYIGLNKNPFRPNRYR"
 BASE COUNT 365 a 303 c 271 g 384 t
 ORIGIN

Query Match 50.3%; Score 1288.8; DB 9; Length 1323;
 Best Local Similarity 99.7%; Pred. No. 6,4e-251;
 Matches 1312; Conservative 2; Mismatches 2; Indels 2; Gaps 2;

QY 1246 AAGCTTTGGAGCAATGTCTACATTCAGAAATCGACCCCATCTGTCTGCAAGGCT 1305
 DB 10 AAGCTTTGGAGCAATGTCTACATTCAGAAATCGACCCCATCTGTCTGCAAGGCT 69
 QY 1306 GCATGATGGGTTTCAGGAGTGTAAAGCTAATGATCATCCGCAAGTCGATGTCTAA 1365
 DB 70 GCATGATGGGTTTCAGGAGTGTAAAGCTAATGATCATCCGCAAGTCGATGTCTAA 129
 QY 1366 TAACTTGACAGAAATAGATGTAGTACACAGGAGCACTTGATCGCATGAAAAACA 1425
 DB 130 TAACTTGACAGAAATAGATGTAGTACACAGGAGCACTTGATCGCATGAAAAACA 189
 QY 1426 GTTGTATCGTATGCAATATGAGGCGACCTCAACACAGAAATGATGACCGCCTCCGCA 1485
 DB 190 GTTGTATCGTATGCAATATGAGGCGACCTCAACACAGAAATGATGACCGCCTCCGCA 249
 QY 1486 CTCGAGAGTGAAGTGGAGAGAGTACGTTCTCAGTGTGACCATGATCTGGCAGATG 1545
 DB 250 CTCGAGAGTGAAGTGGAGAGAGTACGTTCTCAGTGTGACCATGATCTGGCAGATG 309
 QY 1546 GCAAAAGAGTGTCTCTCTGCGACAGAGGAGTGTCTACTCAATTTGAGTGTCTCCACAGTTC 1605
 DB 310 GCAAAAGAGTGTCTCTCTGCGACAGAGGAGTGTCTACTCAATTTGAGTGTCTCCACAGTTC 369
 QY 1606 CCACCTTGTCTGTCTCATACAGCCACACACAGGCTTTGGGACGTATGAACTGTATA 1665
 DB 370 CCACCTTGTCTGTCTCATACAGCCACACACAGGCTTTGGGACGTATGAACTGTATA 429
 QY 1666 ATGACCCGAGGGGCGATACAGCAGAGATGTGACTGTCTTCTTAAGAAATGATGATAT 1725
 DB 430 ATGACCCGAGGGGCGATACAGCAGAGATGTGACTGTCTTCTTAAGAAATGATGATAT 489

OY	380	CCAGTTGGTGTATGACATCTCAGAGAGTTCACCAAAATTCGCCACAAAACCTGGCCGAAC	439
Db	371	CCAGTTGGTGTACACAGAACAGAAATTCACAAACGTCGCCACAAAATTGGACCTCGCTC	430
OY	440	TTTGCTCGCTCGATCTCCACAGTCCCTCCACTGCAGCTACAGTTACAGTTCACCTGCATCTCAAC	499
Db	431	TTTGCTCGCTCCATTTCTTCAGTCATCTACTGCAGAGTTACAGCTACAGCTCAGCGCTTATATAC	490
OY	500	AGATAGCTCTGATGATGAGGTTTCTCCCGAGAAACAGACCAACCACTCCAAAGGCGAC	559
Db	491	AGATAGCTCTGATGATGAGAACATCCGCCAGGAGCAAGACCAAAAGAACTTAAGGGAG	550
OY	560	CAGCAATTTCTGTGTGAABAACATCAAGCAGGAGAAATTTGGACGCCGGGAGATTGATGAT	619
Db	551	CAGTACCTTCTGTGTTAAACATCAACACGCGAGAGATTGGACGAAGAGAAATTTGAAT	610
OY	620	TGCAGAGCAGACATGTCTGTCTGTGATTTCACTCAGAGAAAGTGTCTCAGGGGGAGAAAGCC	679
Db	611	TGCTGACAAAGAAATCCGTCGATTTGATGGCTTTGAGGAAAGAGAGCTCAAGGAGAAAAGCC	670
OY	680	CTTGCGCTGTGCTTAAATAGTGGGCTGTACACATCACAAGCCAGACAGCGGTGTGAT	739
Db	671	TTTGCTGTGAGCCAAATAGTGGTTGCACACACATCAGCTCAGCTGCTGTGCTTAT	730
OY	740	TGAGACACTCTGTGCCCTGGGGGCTCAGTGCAGCTGTGCTGTGAACATCTACTCAAC	799
Db	731	GGAATCTGTGGTGCTCTCTGGGGGCCAGATGCCATGGCGCTGCACACATCTATTCCAC	790
OY	800	TCAGATGAAATGACTGCACAGCACTGGCTGAGGCTGGAATTCAGTCTTCGTTGGAAAGG	859
Db	791	TCTCATGAAAGTGGCGTGCAGCTCTACAGAAAGAGGATTTCTCTTTTGTGCTGGAAGGG	850
OY	860	CGAGTCAGAAAGATGACTTCTGGTGGTGTATTGACCGCTGTGTGAACATGGATGGGTGCA	919
Db	851	AAAGTCAGAGATGACTTTTGGGTGTATCGATAGATGTGTGAATGTGAGGGCTGCGCA	910
OY	920	GGCCAATGATTCCTGATGATGGGGAGAGCTTAACCACTGGGTTTATAGAAGTATCC	979
Db	911	GCCAAATGATGATCTGTGATGATGAGAGGGATCTTACCACATGGATTTATAAAAAGTATCC	970
OY	980	AAACGTGTTTAAAGAAATCCAGACATTTGTGGAAGAGAGCGATGACTGTTCCAGAGCT	1039
Db	971	CAACATGTTTAAAGAAATCAAGGGCATAGTAAAGAGAGATTTACTGGAGTTCAAGAGCT	1030
OY	1040	GATCAGCTCTCCAAAGCTGGGAGGCTGTGTTCGGCCATGAAGCTCAATGATTTCTGT	1099
Db	1031	GTACCAACTGTCCAAAGCTGGGAGGCTGTGTCTCCAGCCATGATGTCAATGACTCACTAGT	1090
OY	1100	TACCAACAGAAATTTGATTAACCTGTACCTGCTGCCAGAAATCCATTTTGATGGCCGTGA	1159
Db	1091	CACCAACAGAAATTTTGAACAACTGTACCTGTGCCGTGAATTCATTTGTATGGACTTAA	1150
OY	1160	GAGGACACAGATGTGATGTTTGGTGGGAAACAAGTGGTGGTGTGGCTATGGTGAAGT	1219
Db	1151	AAGGACACAGACATGATGTTTGGTGGGAAAGCAAGTGTGATGCTGTGGCTATGGAGAGT	1210
OY	1220	AGGCAAGGGCTGTGTGTGCTCTCAAAAGCTTTGGAGCAATTTGTCTACATTACCGAAAT	1279
Db	1211	GGGGAAAGGGTGTGCTGTGCTGCCCTGAAAGCCATGGGCTCCATTGTGTATGTAACCTGAAT	1270
OY	1280	CGACCCCACTGTGTGCTCTCAGGCCCTGCATGATGGGTTCAAGGTTGGTAAAGCTAAATGA	1339
Db	1271	TCACCCCACTGTGTGCCCTCCAAACCTGTAAAGATGGATTCGACTGGTGAATTAATTAAGA	1330
OY	1340	AGTCATCCGGCAAGTGTGATGCTATTAATCTTCACACAGAAATTAAGATGATGAGACAG	1399
Db	1331	GGTCATCCGACAGATGTGACATTTGTATTACCTGTACAGGTACAGAAATGTGGTAAACAG	1390
OY	1400	GGAGCACTTGGATCGCATGAAAAACAAGTTGTATCTGTATGATAATTTGGGCCACTCCAAAC	1459
Db	1391	AAAGCACTTGGACCGTATGAAGAAATAGCTCATCGTTGTATCAATGGGAGCACTTCCAAAC	1450
OY	1460	AGAAATCGATGTACCAAGCTCCGCCACTCCGAGACTGAGCTGGGAGCGAGTACGTTCTCA	1519

Db	1451	AGAGATTGACCTGGCGAGATCTGCGGACACCGAACTGACTGCTGGGAGCGAGTGAATCTTCA	1510
QY	1520	GGTGACCATTGCTCATCTGGCCAGATGGCCAAACGAGTTGTCTCTCGTGGCAGAGGGTCGCT	1579
Db	1511	AGTTGACCATATGTGATATGCGCTCGATNGGCAAGAGATATCTACTGCTGGCAGAGGGCGCGCT	1570
QY	1580	ACTCAATTTGAGCGTCTCCACAGTTCCCGACTTTGTTCGTGCATCACAGGCCAACACACA	1639
Db	1571	GCTGAACTTGTAGCTGCTCCACAGTGCCATACATTTGTGCTGTCAATCACTGCTACTACTCA	1630
QY	1640	GGCTTTGGCACTGATAGAACTCTATATATGCACCCGAGGGCGGATACAAAGCAGATGTGTA	1699
Db	1631	GGCTTCTGCTTGATAGAGCTTTACAAATGCTCTGAGGGGTGGATATAAGCAGGATGTCTA	1690
QY	1700	CTTGCTTCTTAAGAAATGAGTGAATACGTTGCCAGCTTGATGCTGCTGCATATTTGATGC	1759
Db	1691	CCTGTGGCCCAAGAAAGATGAGTGAATGTGGCGCCCTACACCTGCTACTCTTTGAATGC	1750
QY	1760	CGAAGTTACAGAGCGATGAGATGACCAAGCAAAATATCTGGGAGCTCAACAAATATGGGCG	1819
Db	1751	CGACTTGACAGAGCGATGAGATGAAACAGGCGCCAGATGATCTGGGAGCTCAACAAAGAAATGGGCG	1810
QY	1820	ATTCAAACCTAATTTATACAGTACTACTATGAGCCATCTACCAAGGACAGT	1871
Db	1811	CTTCAGGCTTAATTTCTACAGGTATTAAGTTCCTTAACCTCAACACCAATTT	1862

```

RESULT 12
BC024325
LOCUS      BC024325
DEFINITION Homo sapiens, KIAA0828 protein, clone MGC:21525 IMAGE:307552,
ACCESSION  BC024325
VERSION    BC024325.1
KEYWORDS   GI:19353091
SOURCE     MGC.
ORGANISM   Homo sapiens.
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo;
REFERENCE  1 (bases 1 to 5052)
AUTHORS   Strausberg,R.
TITLE     Direct Submission
JOURNAL   Submitted (01-MAR-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK    NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT   MGC help desk

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) Genome DNA Sequencing by: Sequencing Group at the Stanford Human Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: <http://www.simgc.stanford.edu> Contact: (Dickson, Mark) mdc@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov> Series: IRAC Plate: 22 Row: 1 Column: 22 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

```
FEATURES
SOURCE
Location/Qualifiers
1. 5052
/organism="Homo sapiens"
/db_xref="locusid:33382"
/db_xref="taxon:9606"
/clone="MGC:21525 IMAGE:3907552"
/tissue_type="uterus, leiomyosarcoma"
```


DEFINITION Homo sapiens mRNA for KIA0828 protein, partial cds.
 ACCESSION AB020635
 VERSION AB020635.1 GI:4240144
 KEYWORDS
 SOURCE Homo sapiens adult male brain cDNA to mRNA, clone_11b.pbluescriptII SK plus clone:h04230.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Hiroseawa,M., Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
 TITLE Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
 JOURNAL DNA Res. 5 (6), 355-364 (1998)
 MEDLINE 99156230
 REFERENCE 2 (bases 1 to 5025)
 AUTHORS Ohara,O., Suyama,M., Kikuno,R., Nagase,T. and Ishikawa,K.
 TITLE Direct Submission
 JOURNAL Submitted (02-DEC-1998) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology, Yana 1532-3, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913, Fax:+81-438-52-3914)
 FEATURES
 SOURCE
 1..5025
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="h04230"
 /sex="male"
 /tissue_type="brain"
 /clone_11b="pbluescriptII SK plus"
 /dev_stage="adult"
 1..5025
 /gene="KIAA0828"
 <1..1860
 /gene="KIAA0828"
 /codon_start=1
 /product="KIAA0828 protein"
 /protein_id="BAA74851.1"
 /db_xref="GI:4240145"
 /translation="EPVVAEAVMSVOYVSAANAARVPELKDLSPEASESOLGLST
 AAVGMAPAGGDEPAAPAPAEPPVPGSGAPALSPAAKGVQASMKRSDPH
 OHORRDGEALVSPDGYTEAPRYTKTOPADOKEPKPRKIGRRSISRSISOS
 TDYSFSAASYDSDETSPRDKQKNSGSSDFCVKNIKQAFGRREITIAQEMPA
 LALNRKRAOGEKPLGAKIVGCTHTAQTAVLMEVNGMILIDGDLHWYKKYPNF
 AALAESGPPVAMKGESEDFWMCIDRCVNVGEMQNMILIDGDLHWYKKYPNF
 KRIKGIYESVTGVRILQLSKAGKLCVPMANVNDVTOKPDLKCCRESIIDGLR
 TDDMFGKQVYVCGYGVGKGCALAKMGSIVYVTEIDPICALACMDGPRLYKIN
 EVIROVDIVITCTGNKNVYTRHEDLMKNSCIYCNKSHSTEDIVASLRTPELTWERY
 RSQVDHVIWPKRIVLLAEGRLNLSCTVPTFVISTATATQALALIELYNAPEGRY
 KQDYVILPKKMEYVASLHLPFDLHILTELTDQAKYILGNKGPFRKPNTRY"
 BASE COUNT 1261 a 1179 c 1136 g 1429 t
 ORIGIN
 Query Match 38.9%; Score 998; DB 9; Length 5025;
 Best Local Similarity 74.0%; Pred. No. 6.5e-192;
 Matches 1278; Conservative 0; Mismatches 445; Indels 3; Gaps 1;

339 CCTGTACAGCCCGACGAGCACCCTGACCGAGGCGCCGCGACAGTCAGAACATCCAGTT 398
 386 TCGTATGACATGACAGAGTTTACCAAAATTCGCCAACCAAACTGGCCGAAGATCTTTC 445
 399 TCGTACACGAGAGCAGAGATTTAACAAACGTCACCAAAATTTGACGTCCTTTC 458
 446 TCGCTGATCTCAGAGCTCCCTCCAGAGCTACAGCTACAGCTGCTACACAGATAG 505
 459 TCGTTCATTTCTCAGTATCTACTACAGCTACAGCTACAGCTACAGCTACAGTATAG 518
 506 CTCTGATGATGAGGTTTCTCCCGAGAGAGCAGCAAAACCACTCCAGAGGCGACGACAA 555
 519 CTCTGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 578
 566 TTTCTGTGAGAGAACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 635
 579 CTCTGTGTGAGAGAACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 638
 626 GCAAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 685
 639 ACAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 698
 666 TGTGCTAAATATATGAGGCTGACACACATCAGAGAGAGAGAGAGAGAGAGAGAGAG 745
 699 TGGACCCAAATCGTGGGTGACACACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 758
 746 ACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 805
 759 TCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 818
 806 TGAATAGCTGAG 865
 819 TGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 878
 866 AGAAGATGACTTGT 925
 879 AGAAGATGACTTGT 938
 926 CATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 965
 939 CATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 988
 966 GTTTAAGAAATCCGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1045
 999 GTTTAAGAAATCCGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1058
 1046 GCTCTCAAGCTGGAG 1105
 1059 ACTGTCCAAAGCTGGAG 1118
 1106 ACAGAGATTGATTAACCTGTACTGCTGCGAGAGATCCATTTTGATGGCTGAGAGAGAC 1165
 1119 ACAGAGATTGATTAACCTGTACTGCTGCGAGAGATCCATTTTGATGGCTGAGAGAGAC 1178
 1166 CACAGATGATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1225
 1179 AACACACATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1238
 1226 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1285
 1239 AGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1298
 1266 CATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1345
 1299 CATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1358
 1346 CCGGCAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1405
 1359 CCGCAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1418
 1406 CTGTGATGATGAGAAACAGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1465
 1419 CTGTGATGATGAGAAACAGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1478

QY 1466 CGATGTGACCAAGCTCCGACCTCGAGCTGACGTGGGAGAGTGTCTCAGGTGA 1525
 DB 1479 TGAGTGCGCAAGTGTGGGACACCAAGATGACCTGGAGGAGATCTCAAGTTGA 1538
 QY 1526 CCATGTCATCTGGCGAGATGCGCAACAGATGTGTCTCTCTGCGAGAGGTCTCTACTCA 1585
 DB 1539 CCATGTCATCTGGCGAGATGCGCAACAGATGTGTCTCTCTGCGAGAGGTCTCTACTCA 1598
 QY 1586 TTGGAGTGTCTCCACAGTGTCCACCTTTGTCTGTCTCATCAAGCCACACAGAGCTTT 1645
 DB 1599 CCTTAGTGTCTCCACAGTGTCCACCTTTGTCTGTCTCATCAAGCTGTCTACTACAGCTCT 1658
 QY 1646 GGCAGCTGATGAATCTATATATGACCCGAGGAGATACAGAGAGATGTACTGT 1705
 DB 1659 TGCCTGATGAGCTTTACATATGCTCTGAGGCTGCTATTAAGAGATGCTACTGTT 1718
 QY 1706 TCCTAAGAAAATGATGAATACGTTGCGAGCTTGCATCTGCTCATCTTTGATGCCACT 1765
 DB 1719 GCCCAAGAAATGATGAATACGTTGCGAGCTTGCATCTGCTCATCTTTGATGCCACT 1778
 QY 1766 TAAAGAGCTACAGATACCAAGCAAAATATCTGGAGCTCAACAAAATGGCCATTCA 1825
 DB 1779 GACAGAGCTACAGATACCAAGCAAAATATCTGGAGCTCAACAAAATGGCCATTCA 1838
 QY 1826 ACCTAATTATACAGATACCTATGACCAATCTACTACCAAGAGAGCT 1871
 DB 1839 GCTAATATACAGATACCTATGACCTATGACCAATCTACTACCAAGAGAT 1884
 RESULT 14
 AK025372 1987 bp mRNA linear PRI 29-SEP-2000
 LOCUS Homo sapiens cDNA: FLJ21719 fls, clone COLF0094.
 DEFINITION AK025372
 ACCESSION AK025372
 VERSION AK025372.1 GI:10437875
 KEYWORDS o1igo capping; fls (full insert sequence).
 SOURCE Homo sapiens colon mucosa cDNA to mRNA, clone_11b:Colf
 clone:COLF0094.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
 Ota,T., Suzuki,Y., Ohyashi,M., Nishi,T., Shihahara,T., Tanaka,T.,
 Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 1987)
 REFERENCE
 AUTHORS Sugano,S., Suzuki,Y., Ota,T., Odayashi,M., Nishi,T., Isogai,T.,
 Shihahara,T., Tanaka,T. and Nakamura,Y.
 TITLE Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure Analysis, Human
 Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
 Japan (E-mail:cdna@ems.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan, cDNA full insert
 sequencing: Research Association for Biotechnology, cDNA library
 construction, 5'- & 3'-end one pass sequencing: Departent of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).
 FEATURES
 source location/Qualifiers
 1..1987
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="COLF0094"
 /issue_type="colon mucosa"
 /clone_11b="Colf"
 /note="cloning vector pME18SFJ3"
 BASE COUNT 557 a 430 c 501 g 499 t

ORIGIN

Query Match 38.3%; Score 982.2; DB 9; Length 1987;
 Best Local Similarity 78.2%; Pred. No. 9,6e-189;
 Matches 1179; Conservative 0; Mismatches 328; Indels 0; Gaps 0;
 QY 365 GCCAAGAAAGCAATCCAGTTGCTGATGACATCCAGAGTGTACCAATGCCACAA 424
 DB 316 GCGTGAAGTATATCCAGTTGCTGACACAGAGCAAGAAATTCACAAAGCTCCACCA 375
 QY 425 AACTGCCCGAAGATCTTGTGTCTGCTGATCTCAAGTCTCTCACTGACACTAGTTC 484
 DB 376 AATTGAGCGCGCTTTGTCTGCTGCTTCTCAATCTCACTGACACTGACACTG 435
 QY 485 AGCTGATCTTACACAGATAGCTCTGATGATGAGAGTTTCTCCCGAAGAGACCAAC 544
 DB 436 AGCGCTTATATACAGATAGCTCTGATGATGAGACATCCGCCAGGAGCAAGCAAA 495
 QY 545 CAACCTCAAGGAGGAGAGCAATTTCTGTGAGAGAAATCAGAGAGGAGATTTGAGAG 604
 DB 496 GAACCTTAAGGAGAGAGAGTACTTCTGTGTAAGACATCAACAGAGAGATTTGAGAG 555
 QY 605 CCGGAGATTTGAGATTTGACAGAGCAAGATGCTGCTGTGATTTCACTCAGAGAACTGC 664
 DB 556 AAGAGAAATTTGAATTTGCTGACAGAGAAAGCTCTGATGATGAGCTTTGAGAGAGAGC 615
 QY 665 TCAGGGGAGAGAGCCCTTGGCTGCTGCTAAATATGAGGCTGTACACATCAAGCCCA 724
 DB 616 TCAGAGAGAGAGAGCCCTTGGCTGAGCCAAATGCTGGGTTGACACACATCAGCTCA 675
 QY 725 GACAGGAGTGTGATGAGACACTCTGCTGAGGAGCTGAGGCTGCTGCTGCTGCTG 784
 DB 676 GACTGCTGCTTATGAGAACTCTGAGGCTGCTGAGGAGGAGGAGGAGGAGGAGGCTG 735
 QY 785 TAACTACTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 844
 DB 736 CAACATCTATCCACTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 795
 QY 845 GTTCGCTTGAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 904
 DB 796 TTTTGGCTGAG 855
 QY 905 CAGGATGGGTGAG 964
 DB 856 TGGGAGGAGGCTGAG 915
 QY 965 TTTTAAAGATATCCAAAGCTGTTTAAAGAGATCCAGAGCTTTTGAAGAGAGAGCTTAC 1024
 DB 916 TTATAAAAGATATCCAAAGCTGTTTAAAGAGATCCAGAGCTTTTGAAGAGAGAGCTTAC 975
 QY 1025 TGGTGTTCACAGGCTGATAGAGCTCTCCAAAGCTGGAGAGCTGTGTCCGGCATGAA 1084
 DB 976 TGGAGTTCACAGGCTGATAGAGCTCTCCAAAGCTGGAGAGCTGTGTGTCCAGCATGAA 1035
 QY 1085 CGTCAATGATTTCTGTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1144
 DB 1036 TGTCAATGATCTGATACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1095
 QY 1145 TTGGATGGCTTGAAG 1204
 DB 1096 TCTTGTATGAGCTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1155
 QY 1205 TGGCTATGGAGAGTGGAG 1264
 DB 1156 TGGCTATGGAGAGTGGAG 1215
 QY 1265 CTACATTAACGAATCGACCCCATCTGTGCTGACAGGCTGATGATGGGTTCAAGGT 1324
 DB 1216 GATATGATGAAGAAATGAGCCCATCTGTGCTGACAGGCTGATGATGGGTTCAAGGT 1275
 QY 1325 GGTAAAGCTTAATGAAGTCACTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1384
 DB 1276 GGTAAATTAATGAAGTCACTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1335

QY	1848	TEGACCACTACTACCAAGACCAAGTCACCTGTAACCAACACACTCTAAGAATAATTTT	1907
Db	1788	GCATAATATCTGGGAGCTCAACAAATAATGGCCATTCACAACTTAATTTATACAGATAC	1847
Db	61	GCATAATATCTGGGAGCTCAACAAATAATGGCCATTCACAACTTAATTTATACAGATAC	120
QY	1728	GTTCGCACCTGTCATCTCTGCCATCATTTATATGCCACCTTACAGAGCTGACAGATAC	1787
Db	1	GTTCGCACCTGTCATCTCTGCCATCATTTATATGCCACCTTACAGAGCTGACAGATAC	60
QY	1565	GGCAGAGGGTGTCTACTCAATTTAGCTGCTCCACAGTTGCCACCTTTGTTCTGTCCAT	1624
Db	1516	GGCAGAGGGTGTCTACTCAATTTAGCTGCTCCACAGTTGCCACCTTTGTTCTGTCCAT	1575
QY	1655	CACAGCCCAACACAGAGCTTGGCAGCTGATGAATCTATATGACCCGAGGGCGAT	1684
Db	1576	CACAGCCCAACACAGAGCTTGGCAGCTGATGAATCTATATGACCCGAGGGCGAT	1635
QY	1685	CAAGCAGATGTGACTTCTTCTTAAGAAAAATGATGAATACGTTGCCAGCTTCAT	1744
Db	1636	TAAAGCAGATGTGACTTCTTCTTAAGAAAAATGATGAATACGTTGCCAGCTTCAT	1695
QY	1745	GCACATCTTATGATGCCACCTTACAGAGCTGACATGACCAAGCAAAATATCTGGAT	1804
Db	1696	GCACATCTTATGATGCCACCTTACAGAGCTGACATGACCAAGCAAAATATCTGGAT	1755
QY	1805	CAACAAATATGGGCATTCACAACTTAATTTATACAGATCTAATGACCATCTACCAAG	1864
Db	1756	CAACAAATATGGGCATTCACAACTTAATTTATACAGATCTAATGACCATCTACCAAG	1815
QY	1865	GACCACT 1871	
Db	1816	CAGAACT 1822	
RESULT 15			
AR068569			
LOCUS	AR068569	2226 bp	DNA
DEFINITION	Sequence 2 from patent US 5854023.		Linear
ACCESSION	AR068569		
VERSION	AR068569.1	GI:6000776	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1. (bases 1 to 2226)		
TITLE	Hillman, J. L., Corley, N. C., Lai, P. and Shah, P.		
JOURNAL	Polynucleotides encoding human S-adenosyl-5-homocysteine hydrolase		
FEATURES	derived from bladder		
source	Patent: US 5854023-A 2 29-DEC-1998;		
	Location/Qualifiers		
	1..2226		
	/organism="unknown"		
BASE COUNT	649 a 455 c 435 g 675 t.	12 others	
ORIGIN			
Query Match	31.5%; Score 806.2; DB 6; Length 2226;		
Best Local Similarity	99.3%; Pred. No. 4.7e-153;		
Matches 830; Conservative	0; Mismatches 4; Indels 2; Gaps 2;		

Db	121	TGACACATACACCAAGGACCACCGCACCGAACACACACACTCTAAGAAATATTTT	180
Oy	1908	AGATPACTTTATTTTCTTCTTACTCTCTTCCTCTTGATTTTTCCTATPAATTCATTC	1967
Db	181	AGATAACTTTATTTCTTCTTACTCTTCCCTCTGATTTTTCCTPAATTCATTC	240
Oy	1968	TTGTTTTTCACTCATTAATCCAAAGTCGACACACACACAGGAACCTTCCTCAGGCTC	2027
Db	241	TTGTTTTTCACTCATTAATCCAAAGTCGACACACACAGGAACCTTCCTCAGGCTC	300
Oy	2028	TTTATGATGAATATAGATTCAGGCTCCCTCACTCACTATGACTCTAAACAGATTTTACTCC	2087
Db	301	TTTATGATGAATATAGATTCAGGCTCCCTCACTCACTATGACTCTAAACAGATTTTACTCT	360
Oy	2088	CCGACCCACAAAGGATATTCCTCTCTTACCATTTCTGGGACTTATGCTTATTAAG	2147
Db	361	CCGACCCACAAAGGATATTCCT-TCCTTACCATTTCTGGGACTTATGCTTATTAAG	419
Oy	2148	TACCTATATTAACAGAAATGCTAAGGATACCTTCCTGTCGACAACTGCAATGTCTAA	2207
Db	420	TACCTATATTAACAGAAATGCTAAGGATACCTTCCTGTCGACAACTGCAATGTCTAA	479
Oy	2208	TCGCTTTAAAGAGCCCATTTCTTATGCTGTGAATATGATGCTCTTTCATCTTCCAGAG	2267
Db	480	TCGCTTTAAAGAGCCCATTTCTTATGCTGTGAATATGATGCTCTTTCATCTTCCAGAG	539
Oy	2268	AAGCAGGATATGCTACCTACCGGAGGTAGTAAAGTATGGGGGAGCAATGTTAATTC	2327
Db	540	AAGCAGGATATGCTACCTACCGGAGGTAGTAAAGTATGGGGGAGCAATGTTAATTC	599
Oy	2328	CTTAGAAGTTTCCAAAGCCCTGTTTCTGCTGAAGGTGTATGTCCAGTTCCAGAGATGT	2387
Db	600	CTTAGAAGTTTCCAAAGCCCTGTTTCTGCTGAAGGTGTATGTCCAGTTCCAGAGATGT	659
Oy	2388	ATAATGAGCATGGCTTGTATTAAGATCAGAGAGCCCACTGGATTTATATAGTATGCCCTTCC	2447
Db	660	ATAATGAGCATGGCTTGTATTAAGATCAGAGAGCCCACTGGATTTATATAGTATGCCCTTCC	719
Oy	2448	TCCACTCCACACAGACTGCTCATTTTTCAGATTTTAACTAGACATACACTATATTAG	2507
Db	720	TCCACTCCACACAGACTGCTCATTTTTCAGATTTTAACTAGACATACACTGTA-TTGAG	778
Oy	2508	TTTATATTTTGTCTCTAGGATTTATTTCTGTGTGTCAAAAAATTTTTTTTTT	2563
Db	779	TTTATATTTTGTCTCTAGGATTTATTTCTGTGTGTCAAAAAATTTTTTTTTT	834

Search completed: April 20, 2003, 16:28:38
Job time : 6549 secs


```

QY 1728 GTGGCAGCTTGACATCCGATCTGATGATGCCACCTTAGACGCTGACAGATGACCAA 1787
DB 1 GTTCCAGCTTGACATCCGATCTGATGATGCCACCTTAGACGCTGACAGATGACCAA 60
QY 1788 GCAAAATATCTGGGACTCAACAAAATGGCCATTCACAACTTAATTAATAGATACCTAA 1847
DB 61 GCAAAATATCTGGGACTCAACAAAATGGCCATTCACAACTTAATTAATAGATACCTAA 120
QY 1848 TGGACCACTACTACCAAGAGACAGCTGACCTGAAACACACACTCTAAAGAAATATTTTAA 1907
DB 121 TGGACCACTACTACCAAGAGACAGCTGACCTGAAACACACACTCTAAAGAAATATTTTAA 180
QY 1908 AGATACTTTATTTCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1967
DB 181 AGATACTTTATTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
QY 1968 TTGTTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2027
DB 241 TTGTTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY 2028 TTTAGATGAATAGAGAGTTCAGGGTCCCTCACTCTAGTCTACTAAAGAGATTTTACTCC 2087
DB 301 TTTAGATGAATAGAGAGTTCAGGGTCCCTCACTCTAGTCTACTAAAGAGATTTTACTCT 360
QY 2088 CCCAGCCAGAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2147
DB 361 CCCAGCCAGAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 419
QY 2148 TACCTTTATTAACAGAAATGCTAAGGTAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2207
DB 420 TACCTTTATTAACAGAAATGCTAAGGTAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 479
QY 2208 TCGCTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2267
DB 480 TCGCTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539
QY 2268 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2327
DB 540 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
QY 2328 CTTAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2387
DB 600 CTTAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659
QY 2388 ATAATGAGATGAGCTGTAAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2447
DB 660 ATAATGAGATGAGCTGTAAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
QY 2448 TCGACCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2507
DB 720 TCGACCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 778
QY 2508 TTTAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2563
DB 779 TTTAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 834

```

```

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(2226)
OTHER INFORMATION: Polynucleotide encoding human
OTHER INFORMATION: S-adenosyl-5-homocysteine hydrolase (SAHH) derived
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 08/896,005
PATENT FILING DATE: 1997-07-17
PUBLICATION DATE: 1998-12-29
US-09-347-878-3

Query Match      31.5%; Score 806.2; DB 4; Length 2226;
Best Local Similarity 99.3%; Pred. No. 8e-200;
Matches 830; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1728 GTGGCAGCTTGACATCCGATCTGATGATGCCACCTTAGACGCTGACAGATGACCAA 1787
DB 1 GTTCCAGCTTGACATCCGATCTGATGATGCCACCTTAGACGCTGACAGATGACCAA 60
QY 1788 GCAAAATATCTGGGACTCAACAAAATGGCCATTCACAACTTAATTAATAGATACCTAA 1847
DB 61 GCAAAATATCTGGGACTCAACAAAATGGCCATTCACAACTTAATTAATAGATACCTAA 120
QY 1848 TGGACCACTACTACCAAGAGACAGCTGACCTGAAACACACACTCTAAAGAAATATTTTAA 1907
DB 121 TGGACCACTACTACCAAGAGACAGCTGACCTGAAACACACACTCTAAAGAAATATTTTAA 180
QY 1908 AGATACTTTATTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1967
DB 181 AGATACTTTATTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
QY 1968 TTGTTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2027
DB 241 TTGTTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY 2028 TTTAGATGAATAGAGAGTTCAGGGTCCCTCACTCTAGTCTACTAAAGAGATTTTACTCC 2087
DB 301 TTTAGATGAATAGAGAGTTCAGGGTCCCTCACTCTAGTCTACTAAAGAGATTTTACTCT 360
QY 2088 CCCAGCCAGAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2147
DB 600 CCCAGCCAGAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 419
QY 2148 TACCTTTATTAACAGAAATGCTAAGGTAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2207
DB 420 TACCTTTATTAACAGAAATGCTAAGGTAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 479
QY 2208 TCGCTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2267
DB 480 TCGCTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539
QY 2268 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2327
DB 540 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
QY 2328 CTTAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2387
DB 600 CTTAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659
QY 2388 ATAATGAGATGAGCTGTAAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2447
DB 660 ATAATGAGATGAGCTGTAAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
QY 2448 TCGACCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2507
DB 720 TCGACCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 778
QY 2508 TTTAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2563
DB 779 TTTAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 834

```

RESULT 3
US-09-328-111-850/C
Sequence 850, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Aslie, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Dertl, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 850
LENGTH: 636
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(636)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-850

Query Match 15.9%; Score 406.4; DB 4; Length 636;
Best Local Similarity 99.5%; Pred. No. 3.7e-96;
Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

707 TACACATATACAGCCCAAGCAGC-GGTGTTGATTGAGACCTGTGCGGCGGCTC 765
|||||
420 TACACATATACAGCCCAAGCAGCAGCGGTGTTGATTGAGACCTGTGCGGCGGCTC 361
766 AGTGGCGCTGCTGCTGTTGAACATCTACTCACTCAGAAAGAGTAGCTCAGACATGG 825
360 AGTGGCGCTGCTGCTGTTGAACATCTACTCACTCAGAAAGAGTAGCTCAGACATGG 301
826 CTGAGCGCTGAGTTGAGCTGCTGCTGTTGAAAGGCGAGTCAAGAGTAGCTTCTGCTG 885
300 CTGAGCGCTGAGTTGAGCTGCTGCTGTTGAAAGGCGAGTCAAGAGTAGCTTCTGCTG 241
886 GTATTGACCGCTGTGTGAATATGATGGGTGGCAGGCGCAATATATCTGTGATTTGGG 945
240 GTATTGACCGCTGTGTGAATATGATGGGTGGCAGGCGCAATATATCTGTGATTTGGG 181
946 GAACTTAACCACTGGGTTTATTAAGATATCCAAAGCTGTTTAAAGATCCAGGCA 1005
180 GAACTTAACCACTGGGTTTATTAAGATATCCAAAGCTGTTTAAAGATCCAGGCA 121
1006 TTGTGGAAGAGAGCTGACTGTGTTCACAGCTGTATCAGCTCTCCAAAGCTGGGAAC 1065
120 TTGTGGAAGAGAGCTGACTGTGTTCACAGCTGTATCAGCTCTCCAAAGCTGGGAAC 61
1066 TCTGTGTTCCGGCCATGAGCTCAATATTTCTTTACCAACAGAGTTGATACTGT 1125
60 TCTGTGTTCCGGCCATGAGCTCAATATTTCTTTACCAACAGAGTTGATACTGT 1

Patent No. 6210950
GENERAL INFORMATION:
APPLICANT: Johnson, William G.
APPLICANT: Stenroos, Edward S.
TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
FILE REFERENCE: 601-1-057
CURRENT APPLICATION NUMBER: US/09/318,448
CURRENT FILING DATE: 1999-05-25
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 26
LENGTH: 2211
TYPE: DNA
ORGANISM: Homo sapiens
US-09-318-448-26

Query Match 12.3%; Score 314.4; DB 4; Length 2211;
Best Local Similarity 61.9%; Pred. No. 6.1e-72;
Matches 498; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

580 ACATCAGCAGCAGCAATTTGACCGCGGAGATTGAGATTGACAGCAAGACATGCTG 639
79 ACATCGCGCTGCTGCTGCGGAGCGCAAGCCCTGACATTGCTGAGAACAGATGCCG 138
640 CTCTGATTTCACACAGCAAGCTGCTCAGGAGGAGCCCTGGCTGCTAAATAG 699
139 GCGTATGCGGTATGCGGAGCGGTACTCGGCTCCAGCCACTGAGAGGCGCCGCTG 198
700 TGGAGCTGTACACATCACAGCCAGCAGCGGTGTTGATTGAGACACTGTGCGCTG 759
199 CTGCTGCTGCTGACATGACCGGTGAGAGCGCCGCTCATATTGAGACCTGTCAC 258
760 GGGCTCAGTCCGCTGCTGCTGCTGTTGAACATCTACTCACTCAGATGAGTAAGTAC 819
259 GTGCTAGGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 318
820 CACTGCTAGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 879
319 CATTGCCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 378
880 GGTGATGATTTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 939
379 TGTGATGATTTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 438
940 ATGCGGAGAGCTTAACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 999
439 AGCGGAGAGAGCTTAACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 498
1000 GAGGCAATTTGGAAGAGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1059
499 GAGGCAATTTGGAAGAGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 558
1060 GGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1119
559 GGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 618
1120 ACTTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1179
619 ACCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 678
1180 TTGCTGGAAGAGAGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1239
679 TTGCTGGAAGAGAGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 738
1240 CTCTCAAGCTCTTGAAGCAATTTGCTTACATTTACCGAAATGAGCCCATCTGCTGCTG 1299
739 CCTGCGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 798
1300 AGGCTGATGATGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1359
799 AGGCTGATGATGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 858

QY 1360 TCGTAACTTGCACAGGAATA 1383
DB 859 TCTTGTACACACAGAGCTGTA 882

RESULT 5

US-09-347-878-2
Sequence 2, Application US/09347878C
Patent No. 6376210
GENERAL INFORMATION:
APPLICANT: Yuan, Chong
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
FILE REFERENCE: 25885-1651
CURRENT APPLICATION NUMBER: US/09/347,878C
CURRENT FILING DATE: 1999-07-06
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 2211
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Human S-adenosylhomocysteine hydrolase cDNA
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: M61831/genbank
US-09-347-878-2

Query Match 12.3%; Score 314.4; DB 4; Length 2211;
Best Local Similarity 61.9%; Pred. No. 6.1e-72;
Matches 498; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

QY 580 ACATCAGACGAGAAATTGACGCCGGAGATGATGATGACGACCAAGATGCTG 639
DB 79 ACATCGGCTGCTGCTGGGAGCAGAGCCCTGACATGCTGAGAACGAGATGCCG 138
QY 640 CTCTGATTTCACTCAGAGAAAGTGTCTCAGGGGAGAGAGCCCTTGTGCTAAATAG 699
DB 139 GCGTATGGGTATGGGGAGGCGGTACTCGGCTCCAGCCAGCTGAAGGCGCCGATCG 198
QY 700 TGGGCTGTACACATCACACCCAGACGCGGTGTTGATGATGACACTCTGTCCTGG 759
DB 199 CTGGGCTGCTGACATGACCGGTGAGAGCGGCTCTCATGAGACCTGCTCACCTGG 258
QY 760 GGGCTAGAGCGCGCTGTGCTTGTATCATCTACATCACTCAAGTAAGTACTGAG 819
DB 259 GGTCTAGGTGACGTGCTGCTGACGACATCTCTCCACCAACACATGCGCGCTG 318
QY 820 CACTGCTAGAGCTGAGATGACGTGCTTGAAGGCGAGTCAAGATGATCT 879
DB 319 CCATTGCCAAGGTGACATTCGGGTATGCTCTGGAAGGCGAAGACGAGAGTACC 378
QY 880 GGTGTGTATGACCGCTGTGTGAACATGATGAGGTGCAAGCCACATGATCTGATG 939
DB 379 TGTGTGTATGACAGACCTGTACTTCAAGACGCGGCTTCATGATGTTGTGACG 438
QY 940 ATGGGAGACCTAATCCACTGGGTTATAGAAGTATCCAAAGCTTTTAAGAAGTCC 999
DB 439 ACGGGGACACCTCACCACCTATCCACCAAGTACCGGACCTCTGCGACGACATCC 498
QY 1000 GAGGCTATGAGAGAGAGCTGATGTTTCAAGGCTGATACCTCCAAACCTG 1059
DB 499 GAGGCTATGAGAGAGAGCTGATGTTTCAAGGCTGATACCTCCAAACCTG 558
QY 1060 GGAAGCTGTGTGTCGGCCATGACGCTCAATGTTGTATACCAAGAGATTGATA 1119
DB 559 GGAATCTCAAGGCTGCTGCTCAATGTTGTATGTTGTATACCAAGAGATTGATA 618
QY 1120 ACTGTATGCTGCTGAGAGATCCATTTGATGAGCCCTGAAGAGACCAAGATGATG 1179
DB 619 ACTGTATGCTGCTGAGAGATCCATTTGATGAGCCCTGAAGAGACCAAGATGATG 678
QY 1180 TTGGTGGAGAAAGT 1239

DB 679 TTGGCGGACAGGTAGCGGGGTAGCAGGCTATGATGTGGGCAAGGGCTGTGCCAGG 738
QY 1240 CTCTCAAGCTCTTGAGACATTTCTACATTTACGAAATTCAGCCATCTGTGCTG 1299
DB 739 CCTGCGGGGTTTCGAGCCCGCTCATCATCATCAGAGATTGACCCCATCAAGGCTG 798
QY 1300 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
DB 799 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 858
QY 1360 TCGTAACTTGCACAGGAATA 1383
DB 859 TCTTGTACACACAGAGCTGTA 882

RESULT 6

US-09-404-879A-40/c
Sequence 40, Application US/09404879A
Patent No. 6468546
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.462C2
CURRENT APPLICATION NUMBER: US/09/404,879A
CURRENT FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 393
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 40
LENGTH: 292
TYPE: DNA
ORGANISM: Homo sapien
US-09-404-879A-40

Query Match 10.2%; Score 261.4; DB 4; Length 292;
Best Local Similarity 99.3%; Pred. No. 1.2e-58;
Matches 273; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2261 TTCACAGAGCAGGAGTGTACCTACCCGAGAGTGTAGTGTGTGTGTGTGTGT 2320
DB 282 TTCACAGAGCAGGAGTGTACCTACCCGAGAGTGTAGTGTGTGTGTGTGTGT 223
QY 2321 AATTTCCTTGAAGTGTCAAGCCCTGTTCTGCTTAAGGTGTGTGTGTGTGTGT 2380
DB 222 AATTTCCTTGAAGTGTCAAGCCCTGTTCTGCTTAAGGTGTGTGTGTGTGTGT 163
QY 2381 GATGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2440
DB 162 GATGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 103
QY 2441 CCTTCTCTCAGCTCCACAGACTGCTCATTTTTCAGAGTTTAACTAGACTAGCT 2500
DB 102 CCTTCTCTCAGCTCCACAGACTGCTCATTTTTCAGAGTTTAACTAGACTAGCT 43
QY 2501 AATTGATTTAATTTTGTCTCTAGATTTATTC 2535
DB 42 A-TTGAATTTAATTTTGTCTCTAGATTTATTC 9

RESULT 7

US-09-103-840A-1/c
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FRASER, Claire M.
APPLICANT: WHITE, Owen R.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00

; CURRENT APPLICATION NUMBER: US/09/103, 840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patent Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 4411529
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; OTHER INFORMATION: H37Rv
 US-09-103-840A-1

Query Match 9.3%; Score 237.8; DB 4; Length 4411529;
 Best Local Similarity 56.6%; Pred. No. 3e-50;
 Matches 503; Conservative 0; Mismatches 377; Indels 9; Gaps 3;

QY 970 AGAAGTATCCAAACGCTTTAAGAAGATCCGAGCATTGTGGAAGAGCCGTGATGCTG 1029
 DB 3629036 ACAGAGTGAACCAAGATAGCGAGTGGTCAAGAGCGGTCAAGAGAGACACACCGCGG 3628977
 QY 1030 TTCACAGGCTATACGCTCTCCAAAGCTGGAGAGCTCTGTGTCCGGCCATGAAGCTCA 1089
 DB 3628976 TGTGCGGCTCTACCAATTCGCCCGCGGGAATCTGCGCTTCCCGGATCAACGTCA 3628917
 QY 1090 ATGATTCGTTACCAACAAGATTGATTAATCTGCTGCGGAGATCCATTGTTGG 1149
 DB 3628916 ACAGCTCGGTGACCAATTCCTCAATTCGACAAAGTACGCGACTGCGCTCTGATCG 3628857
 QY 1150 ATGCGCTGAAGAGACACAGATGTGATGTTGTGGGAAGAGTGTGTGTGCT 1209
 DB 3628856 AGCGCATCAACCGCGGACCGACCGCTGATCGGGGTGAAGAAGTCCCATCTCGGCT 3628797
 QY 1210 ATGCTAGCTAGGACGAGCGCTGCTGCTCTCAAGCTCTTGAGCAATTTCTACA 1269
 DB 3628796 ACAGCGACGCTCGGTGAGCGCTGCTCGGAGGATGGAAGGCGGAGCGCGCTCTCG 3628737
 QY 1270 TTACGAAATGAGACCCCATCTGCTGCGAGCGCTGATGATGAGGCTGACAGGCTGTA 1329
 DB 3628736 TCACCGATGATGACCCCATCAACCGCTGCGAGCGCTATGAGGCTCTTCAGAGTCTCA 3628677
 QY 1330 AGCTAAATGAAGTCAATCCGCAAGTGTGATGATTAATCTTGACAGGAATTAAGATG 1389
 DB 3628676 CCGTCGAGGAGGCGCATCGGGGACCGCATCTGTAACCGCGACCGGCAACAAAGACA 3628617
 QY 1390 TAGTGACACGGGAGCATTTGATCGCATGAATAAGTGTATCGATGCAATATGGGCC 1449
 DB 3628616 TCATCATGCTCGAGCAATTAAGGCGATGAAGACCAACGCGCATCTCTGGAATATCGGCC 3628557
 QY 1450 ACTCCAAACAGAAATGATGATGACAGCGCTCCGACATCCGAGTGAAGTGGAGGAG 1509
 DB 3628556 ACTTCGAAACAGATGATGATGCGCGGCTGAGAGCGCTCCGAGGCGACACGGGTCAACG 3628497
 QY 1510 TACGTTCTCAGGTGACCATGTCATCTGGCCAGA--TGGCAAAACAGTTGCTCTCTGG 1566
 DB 3628496 TCAAGCTCAGGTGACCATCTGGACCTTTGGGACACAGGGCGCGCTGATCATCTGCTCT 3628437
 QY 1567 CAGAGGCTCTCTACTCAATTTGAGCTGCTCCACAGTTCCAC--TTTGTCTGTCCA 1623
 DB 3628436 CCGAGGGGCGGCTGCTAATCTGGGCAATGCCACCGGCGACCCCTGTCGATGAGCA 3628377
 QY 1624 TCACAGCCAAACAGGCTTTGGACATGATGATGATGATGATGATGATGATGATGATGAT 1683
 DB 3628376 AGAGCTCTGCTAACCAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 3628320
 QY 1684 ACAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1743
 DB 3628319 ACAGCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3628260
 QY 1744 TGCCATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1803
 DB 3628259 TCGAGGCTCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3628200
 QY 1804 TCAACAAAGGCGCATCTCAACCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1852

DB 3628199 TCGAGCTGAGGCTGCTTACAAAGCCGAGCACTACCGTACTGATGCC 3628151

RESULT 8
 US-08-669-536-1
 ; Sequence 1, Application US/08669536
 ; Patent No. 5910444

; GENERAL INFORMATION:
 ; APPLICANT: MASTA, CHIKARA
 ; APPLICANT: UENAKA, KIYOKO
 ; APPLICANT: TANAKA, HIDEO
 ; APPLICANT: KUMETA, SHIGERU

; TITLE OF INVENTION: ORGANISMS IN WHICH THE EXPRESSION OF
 ; TITLE OF INVENTION: S-ADENOSYLHOMOCYSTEINE HYDROLASE GENE IS INHIBITED
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
 ; STREET: PO BOX 747
 ; CITY: FALLS CHURCH
 ; STATE: VA

; COUNTRY: USA
 ; ZIP: 22040-0747
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/669,536
 ; FILING DATE:
 ; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURPHY JR, GERALD M
 ; REGISTRATION NUMBER: 28,977
 ; REFERENCE/DOCKET NUMBER: 1254-128
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 205-8000
 ; TELEFAX: (703) 205-8050

; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1812 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA

US-08-669-536-1

Query Match 8.9%; Score 227.4; DB 2; Length 1812;
 Best Local Similarity 56.1%; Pred. No. 2.3e-49;
 Matches 514; Conservative 0; Mismatches 391; Indels 12; Gaps 4;

QY 972 AGTATCCAAACGCTTTAAGAAGATCCGAGCATTGTGGAAGAGCCTGACTGCTGTT 1031
 DB 634 AATATATCCAAAGATGAAGAAAGACTGCTGCTGTTGAGAGAACTACATGAGT 693
 QY 1032 CACAGGCTGTACAGCTTCCAAAGCTGGGAAGCTGTGTCTCGGCCATGAAGTCAAT 1091
 DB 694 AAGAGGCTTTATCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 753
 QY 1092 GATCTGTACCAACAGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1151
 DB 754 GATCTGTACCAACAGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 813
 QY 1152 GCGCTGAAGAGACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1211
 DB 814 GGTCTATGAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 873
 QY 1212 GGTAGGTGAGCAAGGCTGCTGCTCTCTCAAGCTCTTGAGACATTTGTCATAT 1271
 DB 874 GGAATGTCGCAAGGCTGCTGCTCTCTCTCAAGCTCTTGAGACATTTGTCATAT 933
 QY 1272 ACCGAATTCAGCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1331

Db 934 ACCGATTTGACCTTATCTGCTCCAGGCTACATGGAAGGCTCCAGTCTTACT 993
 Qy 1332 CTAAATGAAGTATCCGGGAGTGGATGCTATTAATCTGCACAGAAATGAATGA 1391
 Db 994 CTAGAGAGTGTCTTCTGATGTTGATCTTGTCTACACAGCGGATGAACAGACTT 1053
 Qy 1392 GTGACACGGGAGACTTGTGATCCGATGAAAAACAGTGTGATCTGATGAATGGCCAC 1451
 Db 1054 ATCATGTTGATACCAATGAGGAAGATGAAGAACATGCCATGTTGTCACATTTGGTAC 1113
 Qy 1452 TCCACACAGAAATGATGATGACACGCTCCGAC---TCCGAGAGTACCTGGAGGCA 1508
 Db 1114 TTGACACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173
 Qy 1509 GTACGTTCTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1565
 Db 1174 ATTAAGGCTCAAAACGACAGATGGGATGCTCCCTGACACACAGATGATGATGATGATG 1233
 Qy 1566 GCAGAGGCTGCTACCAATTTGAGCTGCTCCACAG---TTCCACCTTTGTTCTGTC 1622
 Db 1234 GCTGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1293
 Qy 1623 ATGACAGCCCAACACAGGCTTTGGCACTGATGATGATGATGATGATGATGATGATGAT 1679
 Db 1294 TGCTCTTCTACCTACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1353
 Qy 1680 CGATACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1739
 Db 1354 AAGTATGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1413
 Qy 1740 CACTGCTCATCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799
 Db 1414 CACTGCTGAGAGCTGCGAGCGGACCTTACCAACTTTCGAGAGATGAGAGATGATGATGAT 1473
 Qy 1800 GGAATCAAAATGAGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1859
 Db 1474 AGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1533
 Qy 1860 CCAGAGCAGCTCCACC 1876
 Db 1534 CGACAGAGAGAGAGAGC 1550

RESULT 9

US-08-930-894-1
 Sequence 1, Application US/08930894
 Patent No. 6037524

GENERAL INFORMATION:

APPLICANT: GREENLAND, Andrew James
 APPLICANT: DRAVER, John
 APPLICANT: SKIPSEY, Marc
 APPLICANT: WANNER, Simon
 TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEIN HYDROLASE PROMOTER
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pillsbury Madison & Sutto
 STREET: 1100 New York Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005-3918

COMPUTER-READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: MS Word
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/930,894
 FILING DATE: 09-OCT-1997
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/GB96/00882
 FILING DATE: 10-APR-1996

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9507381.3
 FILING DATE: 10-APR-1995
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1767 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: SHH GENE FROM ASPARAGUS
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 26..1483
 OTHER INFORMATION: /codon_start= 26
 US-08-930-894-1

Query Match 8.5%; Score 216.6; DB 3; Length 1767;
 Best Local Similarity 55.9%; Pred. No. 1.5e-46;
 Matches .496; Conservative 0; Mismatches 379; Indels 12; Gaps 4;

Qy 972 AAGTATCAAGGTGTTTAAAGATCCAGGCTTGTGGAAGAGCGTGCATGTGT 1031
 Db 596 AAGTATCAAGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 655
 Qy 1032 CACAGGCTGATACGCTCTCCAAAGCTGAGAGCTGTGTTCCGGCAATGATGATGAT 1091
 Db 656 AAGAGCTTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 715
 Qy 1092 GATTCTGTTACCAAGAGATTTGATTAATGTTGATCTGCTGCTGCTGCTGCTGCTGCT 1151
 Db 716 GACTCCGTCACCAAGAGAGATTTGATTAATGTTGATGATGATGATGATGATGATGATGAT 775
 Qy 1152 GGCTCAAGAGGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1211
 Db 776 GGTCTATGAGGAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 835
 Qy 1212 GGTGATGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1271
 Db 836 GGTGATGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895
 Qy 1272 ACCGAATGACCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1331
 Db 896 ACGGATGATGACCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955
 Qy 1332 CTAAATGAAGTATCCGGGCAAGTGCATGCTGATTAATCTGCAAGAGAAATGAATGA 1391
 Db 956 CTGAGAGATGTTCTCAGAGGCGGATATCTTTGTTACCAACCGGATGAACAGAGATC 1015
 Qy 1392 GTGACAGGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1451
 Db 1016 ATCATCTGAGACCATGAGAGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 1075
 Qy 1452 TCCACACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1508
 Db 1076 TTGACACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1135
 Qy 1509 GTACGTTCTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1565
 Db 1136 ATCAAGCCCCAGACTGACCGGTGCTTCCCTGAACCAAGCACTGATATATGTTT 1195
 Qy 1566 GCAGAGGCTGCTACCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1622
 Db 1196 GCTGAGGCGGATCATGAACCTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1255
 Qy 1623 ATCAAGCCCAACAGAGGCTTTGGCACTGATGATGATGATGATGATGATGATGATGATGAT 1679
 Db 1256 TGCTCTTCAACCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1315
 Qy 1680 CGATACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739

RESULT 12
US-09-081-395-8
; Sequence 8, Application US/09081395
; Patent No. 6083746
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazon, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive, 32nd Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,395
FILING DATE: 18-May-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 6083746nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-KK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-081-395-8
Query Match 2.88; Score 72.6; DB 3; Length 289;
Best Local Similarity 58.68; Pred. No. 1.5e-09;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
DB 580 ACATCAGCAGGAGAAATTGGAGCGCGGAGATTGAGATTGCAGAGCAATGCTG 639
DB 75 ACATCGGCTGGCTGCTGGGAGGAGCGAGCCCTGACATTCCTGAGAGAGATGCCG 134
QY 640 CTCGATTTCACCTAGGAAAGCTGCTCAGGGGAGAGAGCCCTGGCTGTAAATAG 699
DB 135 GCGTATGCGTATCGGGAGGAGGACTGCGGCTCCAGGCACTGAAGGCGCCGATG 194
QY 700 TGGGCTGTACACATCATCAGCCAGACAGCGGTGTATTAGACACTCTGCCCCG 759
DB 195 CTGGCTGCTGACATGACCTGAGAGCGGCGTCTCTATTGAGACCTGACACCTG 254
QY 760 GGGCTCAGTGGCGCTGCTGCTGTATTGATCATCTAC 794
DB 255 GTGCTGAGGTGAGTGTGCTGACGCTGCAACATCTTC 289

APPLICANT: Kazarov, Alexander
APPLICANT: Mazon, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allgretti & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/416,833
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/204,740
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 6197521nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-416-833-8
Query Match 2.88; Score 72.6; DB 4; Length 289;
Best Local Similarity 58.68; Pred. No. 1.5e-09;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
DB 580 ACATCAGCAGGAGAAATTGGAGCGCGGAGATTGAGATTGCAGAGCAATGCTG 639
DB 75 ACATCGGCTGGCTGCTGGGAGGAGCGAGCCCTGACATTCCTGAGAGAGATGCCG 134
QY 640 CTCGATTTCACCTAGGAAAGCTGCTCAGGGGAGAGAGCCCTGGCTGTAAATAG 699
DB 135 GCGTATGCGTATCGGGAGGAGGACTGCGGCTCCAGGCACTGAAGGCGCCGATG 194
QY 700 TGGGCTGTACACATCATCAGCCAGACAGCGGTGTATTAGACACTCTGCCCCG 759
DB 195 CTGGCTGCTGACATGACCTGAGAGCGGCGTCTCTATTGAGACCTGACACCTG 254
QY 760 GGGCTCAGTGGCGCTGCTGCTGTATTGATCATCTAC 794
DB 255 GTGCTGAGGTGAGTGTGCTGACGCTGCAACATCTTC 289

RESULT 14
PCT-US95-02521-8
; Sequence 8, Application PC/TUS9502521
; GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/02521

FILING DATE:

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 289 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

PCT-US95-02521-8

Query Match 2.8%; Score 72.6; DB 5; Length 289;

Best Local Similarity 58.6%; Pred. No. 1.5e-09; Mismatches 89; Indels 0; Gaps 0;

Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 580 ACATCAGCAGCAGCAATTTGGACGCCGGAGATTGAGATTGAGACCAAGACATGCTG 639

DB 75 ACATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 134

QY 640 CTCTATTTCACCTAGAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 699

DB 135 GCTCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 194

QY 700 TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 759

DB 195 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 254

QY 760 GGGCTCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 794

DB 255 GTGCTGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 289

QY 760 GGGCTCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 794

DB 255 GTGCTGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 289

RESULT 15

US-08-204-740-6

Sequence 6, Application US/08204740

Patent No. 5753432

GENERAL INFORMATION:

APPLICANT: Gudkov, Andrei

APPLICANT: Kazarov, Alexander

APPLICANT: Mazo, Ilya

APPLICANT: Roninson, Igor B

TITLE OF INVENTION: Methods for Identifying Genetic

TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allgretti & Mitcoff, Ltd.

STREET: 10 S. Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/204,740

FILING DATE: 04-MAR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5753432man, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 93,354-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 285 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-204-740-6

Query Match 2.7%; Score 68.4; DB 1; Length 285;

Best Local Similarity 56.8%; Pred. No. 1.9e-08; Mismatches 96; Indels 0; Gaps 0;

Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 573 GTGAAGACATCAAGCAGCAGCAATTTGGACGCCGGAGATTGAGATTGAGACCAAGAC 632

DB 64 GTGCGGACATGGAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 123

QY 633 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 692

DB 124 ATGCCAGGCTTGTATGCGCATGCGGAGATGTACTCAGCCCTCAGCCCTGAGAGGCTGCT 183

QY 693 AAAATAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 752

DB 184 CGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243

QY 753 GCCCTGGGGCTCAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 794

DB 244 GCCCTGGGGCTCAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285

Search completed: April 20, 2003, 18:17:48
Job time : 3273 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 18:50:19 ; Search time 123 Seconds

(without alignments)
6390.346 Million cell updates/sec

Title: US-09-782-051-1

Perfect score: 2563
Sequence: 1 ggcgcggcagctcgagct.....aaaaaaaaaaaaaaaaaaaaa 2563Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents, NA:*

- 1: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/lna/5C_COMB.seq:*
- 4: /cgn2_6/ptodata/1/lna/5D_COMB.seq:*
- 5: /cgn2_6/ptodata/1/lna/5E_COMB.seq:*
- 6: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	553	21.6	2226	2 US-08-896-005-2	Sequence 2, Appl1
2	553	21.6	2226	4 US-09-347-878-3	Sequence 3, Appl1
3	370	14.4	636	4 US-09-328-111-850	Sequence 850, App
4	237	9.2	232	4 US-09-404-879A-40	Sequence 40, Appl
5	26	1.0	787	1 US-08-236-427-12	Sequence 12, Appl
6	25	1.0	632	4 US-09-328-111-798	Sequence 12, Appl
7	25	1.0	685	4 US-09-227-357-66	Sequence 66, Appl
8	25	1.0	989	4 US-09-345-882-3	Sequence 3, Appl1
9	25	1.0	2933	4 US-09-262-773-209	Sequence 209, App
10	25	1.0	23071	4 US-09-262-773-210	Sequence 210, App
11	25	1.0	162450	4 US-09-345-882-1	Sequence 1, Appl1
12	23	0.9	56	2 US-07-920-281C-25	Sequence 25, Appl
13	23	0.9	80	4 US-08-466-277-25	Sequence 25, Appl
14	23	0.9	256	4 US-09-385-982-187	Sequence 187, App
15	23	0.9	322	4 US-09-385-982-216	Sequence 216, App
16	23	0.9	322	4 US-09-385-982-216	Sequence 216, App
17	23	0.9	322	4 US-09-385-982-216	Sequence 216, App
18	23	0.9	322	4 US-09-385-982-216	Sequence 216, App
19	23	0.9	322	4 US-09-385-982-216	Sequence 216, App
20	23	0.9	452	4 US-09-397-787-237	Sequence 237, App
21	23	0.9	479	4 US-09-328-111-432	Sequence 432, App
22	23	0.9	506	4 US-09-149-476-252	Sequence 252, App
23	23	0.9	550	4 US-08-998-416-148	Sequence 148, App
24	23	0.9	571	1 US-08-322-742-14	Sequence 14, Appl
25	23	0.9	580	2 US-08-809-763-2	Sequence 2, Appl1
26	23	0.9	580	3 US-08-956-253-2	Sequence 2, Appl1
27	23	0.9	603	2 US-08-924-838-6	Sequence 6, Appl1

c 28	23	0.9	608	4 US-09-385-982-523	Sequence 523, App
c 29	23	0.9	622	4 US-09-385-982-312	Sequence 312, App
c 30	23	0.9	634	4 US-09-385-982-359	Sequence 359, App
c 31	23	0.9	658	4 US-09-328-111-816	Sequence 816, App
c 32	23	0.9	660	4 US-09-643-597-96	Sequence 96, Appl
c 33	23	0.9	669	4 US-09-328-111-782	Sequence 782, App
c 34	23	0.9	688	6 5498694-3	Patent No. 5498694
c 35	23	0.9	718	4 US-09-123-912-96	Sequence 96, Appl
c 36	23	0.9	719	4 US-09-227-357-74	Sequence 74, Appl
c 37	23	0.9	735	3 US-08-950-720A-5	Sequence 3, Appl1
c 38	23	0.9	826	1 US-08-698-551-3	Sequence 3, Appl1
c 39	23	0.9	826	2 US-08-602-228-3	Sequence 3, Appl1
c 40	23	0.9	826	2 US-08-649-341A-3	Sequence 3, Appl1
c 41	23	0.9	826	2 US-08-494-440B-3	Sequence 3, Appl1
c 42	23	0.9	826	2 US-08-533-901B-3	Sequence 3, Appl1
c 43	23	0.9	826	2 US-08-839-031A-3	Sequence 3, Appl1
c 44	23	0.9	826	2 US-08-839-031A-3	Sequence 3, Appl1
c 45	23	0.9	826	4 US-09-185-258C-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-896-005-2
; Sequence 2, Application US/08896005
; Patent No. 5854023
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purni
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,005
; FILING DATE: Filed Herewith
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2226 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADTUT04
; CLONE: 1519044
; US-08-896-005-2

Query Match 21.6%; Score 553; DB 2; Length 2226;
Best Local Similarity 99.6%; Pred. No. 1.4e-236;
Matches 773; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY	1728	GTGGCAGCTTCATCTGGCATCATTTGATGAGCCACCTTACAGAGCTGACAGATGACC	1787
Db	1	GTGGCAGCTTCATCTGGCATCATTTGATGAGCCACCTTACAGAGCTGACAGATGACC	60
QY	1788	GCAAAATATCTGGGACTCAACAAAAATGGGCAATTCGAAACCTATTTATACAGATAC	1847
Db	61	GCAAAATATCTGGGACTCAACAAAAATGGGCAATTCGAAACCTATTTATACAGATAC	120
QY	1848	TGGACCATACCTACCAAGGACAGTCCACTGAAACACACACTCTAAAGAAATATTTT	1907
Db	121	TGGACCATACCTACCAAGGACAGTCCACTGAAACACACACTCTAAAGAAATATTTT	180
QY	1908	AGATACCTTTATTTTCTCTTACTCTCTTCCCTCTGATTTTTCCTATATTTCAATC	1967
Db	181	AGATACCTTTATTTTCTCTTACTCTCTTCCCTCTGATTTTTCCTATATTTCAATC	240
QY	1968	TTGTTTTTCACTCTATTATCCAGATTCTGCAACACACAGGAACTTGCTTACGGCTC	2027
Db	241	TTGTTTTTCACTCTATTATCCAGATTCTGCAACACACAGGAACTTGCTTACGGCTC	300
QY	2028	TTTATATGAATAGAAAGTCAAGGTCGCCCTCACCTCTGCTACTAAAGAAAGTTTACTCC	2087
Db	301	TTTATATGAATAGAAAGTCAAGGTCGCCCTCACCTCTGCTACTAAAGAAAGTTTACTCT	360
QY	2088	CCCAAGCCCAAGAAAGTGATTTCTTCTTACCACTTTCTGGGACTTTAGTCTTATTAAG	2147
Db	361	CCCAAGCCCAAGAAAGTGATTTCTTCTTACCACTTTCTGGGACTTTAGTCTTATTAAG	419
QY	2148	TACCTTATTAACAGAAATGCTAAGGTACCTTCTGTGTGGAACAATCTGCANCTTAA	2207
Db	420	TACCTTATTAACAGAAATGCTAAGGTACCTTCTGTGTGGAACAATCTGCANCTTAA	479
QY	2208	TGCGCTTAAAGAGCCACTTCTTACTGCTGTAATCAGTCTTCACTTCTTCAGAG	2267
Db	480	TGCGCTTAAAGAGCCACTTCTTACTGCTGTAATCAGTCTTCACTTCTTCAGAG	539
QY	2268	AAGCAGGATGTACTACCTACCCGGCAGGTAGATTAGATGTGGTGTGCATGTATTTCC	2327
Db	540	AAGCAGGATGTACTACCTACCCGGCAGGTAGATTAGATGTGGTGTGCATGTATTTCC	599
QY	2328	CTTAAGATCTCAAGAGCCCTGTTTCCCTGCGTAAAGGTGTATGTCCACTTACGAATGTGT	2387
Db	600	CTTAAGATCTCAAGAGCCCTGTTTCCCTGCGTAAAGGTGTATGTCCACTTACGAATGTGT	659
QY	2388	ATAATGAGATGGCTGTGAATGATCAGAGGCCACTTGATTTATGTATAGCCCTCC	2447
Db	660	ATAATGAGATGGCTGTGTGAATGATCAGAGGCCACTTGATTTATGTATAGCCCTCC	719
QY	2448	TCCACTCCCAACAGACTTGCTCATTTTTCGAGTCTTTTAACTAGACTACACTCTATT	2503
Db	720	TCCACTCCCAACAGACTTGCTCATTTTTCGAGTCTTTTAACTAGACTACACTCTATT	775
RESULT 2			
US-09-347-878-3			
Sequence 3, Application US/09347878C			
Patent No. 6376210			
GENERAL INFORMATION:			
APPLICANT: Yuan, Chong			
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES			
FILE REFERENCE: 25885-1651			
CURRENT APPLICATION NUMBER: US/09/347,878C			
CURRENT FILING DATE: 1999-07-06			
NUMBER OF SEQ ID NOS: 75			
SOFTWARE: PatentIn Ver. 2.0			
SEQ ID NO 3			
LENGTH: 2226			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
NAME/KEY: misc_feature			
LOCATION: (1)..(2226)			

; OTHER INFORMATION: Polynucleotide encoding human
 ; OTHER INFORMATION: S-adenosyl-5-homocysteine hydrolase (SAHH) derived
 ; OTHER INFORMATION: from bladder; n=a, c, g, or t
 ; PUBLICATION INFORMATION:
 ; PATENT DOCUMENT NUMBER: 08/896,005
 ; PATENT FILING DATE: 1997-07-17
 ; PUBLICATION DATE: 1998-12-29
 US-09-347-878-3

Query Match	21.6%	Score 553;	DB 4;	length 2226;
Best Local Similarity	99.6%	Pred. No. 1,4e-236;		
Matches 773;	Conservative	0;	Mismatches	1;
			Indels	1;
			Gaps	1;

OY	1728	GTGGCAGCTTCATCTGGCATCATTTGATGCGCCACTTACAGAGCTGACAGATGACC	1787
Db	1	GTGGCAGCTTCATCTGGCATCATTTGATGCGCCACTTACAGAGCTGACAGATGACC	60
OY	1788	GCAAAATATCTGGGACTCACAACAAAATGGGCCATTCAACCTAATTAATACAGATCTAA	1847
Db	61	GCAAAATATCTGGGACTCACAACAAAATGGGCCATTCAACCTAATTAATACAGATCTAA	120
OY	1848	TGGACCAATCTACCAAGGACAGCTCACCCTGAAACACACACTTAAGAAATATTTT	1907
Db	121	TGGACCAATCTACCAAGGACAGCTCACCCTGAAACACACACTTAAGAAATATTTT	180
OY	1908	AGATACCTTTTATTTCTCTACTCTCTCTCCCTGATTTTTCCTAATATTCATTC	1967
Db	181	AGATACCTTTTATTTCTCTACTCTCTCTCCCTGATTTTTCCTAATATTCATTC	240
OY	1968	TTGTTTTTCATCTCATTAATGCCAAGTTCTGCAGACACACAGGAACTTGCTATGGCTC	2027
Db	241	TTGTTTTTCATCTCATTAATGCCAAGTTCTGCAGACACACAGGAACTTGCTATGGCTC	300
OY	2028	TTTAATATGAAATAGAAGTGCAGGTCCTCCATCTCTAGTCACTAAAGAAATTTTACTCC	2087
Db	301	TTTAATATGAAATAGAAGTGCAGGTCCTCCATCTCTAGTCACTAAAGAAATTTTACTCT	360
OY	2088	CCGACCCCAAGAGGATCTTCTCTTACATCTTCGCGGAGCTTACTGATTAATAGG	2147
Db	361	CCGACCCCAAGAGGATCTTCTCTTACATCTTCGCGGAGCTTACTGATTAATAGG	419
OY	2148	TACCTTAATTAACAGGAAATGCTAAGGTACTCTCTGTGGACAAATCTGCAATCTTAA	2207
Db	420	TACCTTAATTAACAGGAAATGCTAAGGTACTCTCTGTGGACAAATCTGCAATCTTAA	479
OY	2208	TGCGCTTAAAGAGCCCATTTCTTAGCTGCTGAATACAGTCTTCACTTCTTCAGAG	2267
Db	480	TGCGCTTAAAGAGCCCATTTCTTAGCTGCTGAATACAGTCTTCACTTCTTCAGAG	539
OY	2268	AAGCAGGGATGTACTCTACCGGACAGTGAAGTATGATGGTGGTGCATGTAATTTCC	2327
Db	540	AAGCAGGGATGTACTCTACCGGACAGTGAAGTATGATGGTGGTGCATGTAATTTCC	599
OY	2328	CTTAGAAGTCCAAAGCCCTGTTTCCCTGCGTAAAGGTGATATGTCACAGTTCAGAGATGT	2387
Db	600	CTTAGAAGTCCAAAGCCCTGTTTCCCTGCGTAAAGGTGATATGTCAGAGATGTGT	659
OY	2388	ATAATGAGATGGCTTGTTAAGATCAGAGAGCCCACTTGATTTATATGTAAGCCCTTC	2447
Db	660	ATAATGAGATGGCTTGTTAAGATCAGAGAGCCCACTTGATTTATATGTAAGCCCTTC	719
OY	2448	TCCATCTCCACAGACTTCTCATTTTTCAGAGTTTAACTATACATCACTCAATTT	2503
Db	720	TCCATCTCCACAGACTTCTCATTTTTCAGAGTTTAACTATACATCACTCAATTT	775

RESULT3
US-09-328-111-850/c
; Sequence 850, Application US/09328111
; Patent No. 626233
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.


```

; APPLICANT: Attle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 850
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(636)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-850

```

```

Query Match          14.4%; Score 370; DB 4; Length 636;
Best Local Similarity 100.0%; Pred. No. 3.4e-155;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 756 CTGGGGGCTGATGCGCGCTGCTGCTGTAACATCTACTCACTCAGATGAGTAGCT 815
DB 370 CTGGGGGCTGATGCGCGCTGCTGCTGTAACATCTACTCACTCAGATGAGTAGCT 311
QY 816 GCAGCACTGCTGAGCGCTGAGTGCAGTGTTCCTTGGAAGGGGAGTCAGAAATGAC 875
DB 310 GCAGCACTGCTGAGCGCTGAGTGCAGTGTTCCTTGGAAGGGGAGTCAGAAATGAC 251
QY 876 TTCTGGGCTGATGCGCGCTGCTGCTGTAACATCTACTCACTCAGATGAGTAGCT 935
DB 250 TTCTGGGCTGATGCGCGCTGCTGCTGTAACATCTACTCACTCAGATGAGTAGCT 191
QY 936 GATGATGAGGAGAGCTTAACCCACTGAGTGTATAGAAGATCCAAAGCTGTTAAGAG 995
DB 190 GATGATGAGGAGAGCTTAACCCACTGAGTGTATAGAAGATCCAAAGCTGTTAAGAG 131
QY 996 ATCCGAGGCTTTGTGAAGAGAGCGTGAATGCTGTTACAGGCTGTATGCTTCCAAA 1055
DB 130 ATCCGAGGCTTTGTGAAGAGAGCGTGAATGCTGTTACAGGCTGTATGCTTCCAAA 71
QY 1056 GCTGGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1115
DB 70 GCTGGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11
QY 1116 GATACTGT 1125
DB 10 GATACTGT 1

```

```

; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 292
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-40

Query Match          9.2%; Score 237; DB 4; Length 292;
Best Local Similarity 100.0%; Pred. No. 4.8e-96;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2267 GAAGCAGGAGTGTACTACCCCGGAGGTAGATGATGAGTGTGCTGATTAATTC 2326
DB 276 GAAGCAGGAGTGTACTACCCCGGAGGTAGATGATGAGTGTGCTGATTAATTC 217
QY 2327 CCTAGAGATTCAGAGCCCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2386
DB 216 CCTAGAGATTCAGAGCCCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 157
QY 2387 TATATGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2446
DB 156 TATATGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 97
QY 2447 CTCACCTCCACACAGACTGCTCATTTTTCGAGTTTACTAGACTACACTCTAT 2503
DB 96 CTCACCTCCACACAGACTGCTCATTTTTCGAGTTTACTAGACTACACTCTAT 40

RESULT 5
US-08-236-427-12
; Sequence 12, Application US/08236427
; Patent No. 5541310
; GENERAL INFORMATION:
; APPLICANT: Ward, Eric
; APPLICANT: Volzath, Sandra
; APPLICANT: Kotsuul, Sain-ichi
; APPLICANT: Tada, Sachio
; APPLICANT: Mori, Ichiro
; APPLICANT: Iwasaki, Genji
; TITLE OF INVENTION: Herbicide Resistant Plants
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/236,427
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/061,644
; FILING DATE: 13-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: S-19558/P1/CGC 1706/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8615
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 787 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: pIGPdat.2
US-08-236-427-12

Query Match 1.0%; Score 26; DB 1; Length 787;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2538 TTGTCTCAAAAAAAAAAAAAAAAA 2563
|||||
DB 761 TTGTCTCAAAAAAAAAAAAAAAAA 786

RESULT 6
US-09-328-111-798/c

Sequence 798, Application US/09328111
Patent No. 6262333

GENERAL INFORMATION:

APPLICANT: Endege, Wilson O.

APPLICANT: Steinmann, Kathleen E.

APPLICANT: Astle, Jon H.

APPLICANT: Burgess, Christopher C.

APPLICANT: Bushnell, Steven E.

APPLICANT: Catroli III, Eddie

APPLICANT: Catino, Theodore J.

APPLICANT: Dertl, Adnan

APPLICANT: Ford, Donna M.

APPLICANT: Lewis, Marcia E.

APPLICANT: Monahan, John E.

APPLICANT: Schlegel, Robert

TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

FILE REFERENCE: CCD-257 (US)

CURRENT APPLICATION NUMBER: US/09/328,111

CURRENT FILING DATE: 1999-06-08

EARLIER APPLICATION NUMBER: US 60/088,801

EARLIER FILING DATE: 1998-06-10

NUMBER OF SEQ ID NOS: 850

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 798

LENGTH: 632

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)..(632)

OTHER INFORMATION: n - A,T,C or G

US-09-328-111-798

Query Match 1.0%; Score 25; DB 4; Length 632;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2539 TGTGCAAAAAAAAAAAAAAAAAA 2563
|||||
DB 40 TGTGCAAAAAAAAAAAAAAAAAA 16

RESULT 7

US-09-227-357-66/c

Sequence 66, Application US/09227357

Patent No. 6342581

GENERAL INFORMATION:

APPLICANT: Fischer et al.

TITLE OF INVENTION: 123 Human Secreted Proteins

FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/227,357
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 66
LENGTH: 685

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (7)
OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-66

Query Match
Best Local Similarity 100.0%; Score 25; DB 4; Length 685;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2539 TGTCCAAAAA
DB 78 TGTCCAAAAA 54

RESULT 8
US-09-345-882-3/C
Sequence 3, Application US/09345882
Patent No. 6399373
GENERAL INFORMATION:
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
FILE REFERENCE: GENSET.031A
CURRENT APPLICATION NUMBER: US/09/345,882
CURRENT FILING DATE: 1999-06-30
PRIOR APPLICATION NUMBER: US 60/091,315
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/111,909
PRIOR FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 140
SOFTWARE: Patent.pm
SEQ ID NO 3
LENGTH: 999
TYPE: DNA
ORGANISM: Homo sapiens
US-09-345-882-3

Query Match
Best Local Similarity 100.0%; Score 25; DB 4; Length 999;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2539 TGTCCAAAAA
DB 208 TGTCCAAAAA 184

RESULT 9
US-09-262-773-209
Sequence 209, Application US/09262773
Patent No. 6225451
GENERAL INFORMATION:
APPLICANT: Ballinger, Dennis G.
APPLICANT: Ding, Wei
APPLICANT: Wagner, Susanne
APPLICANT: Hess, Mark A.
TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
FILE REFERENCE: Myriad 3
CURRENT APPLICATION NUMBER: US/09/262,773
CURRENT FILING DATE: 1999-03-04
NUMBER OF SEQ ID NOS: 210
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 209
LENGTH: 2933
TYPE: DNA
ORGANISM: human
US-09-262-773-209

Query Match
Best Local Similarity 100.0%; Score 25; DB 4; Length 2933;
Matches 100.0%; Pred. No. 0.079;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2539 TGTCCAAAAA 2563
DB 920 TGTCCAAAAA 944

RESULT 10
US-09-262-773-210
Sequence 210, Application US/09262773
Patent No. 6225451
GENERAL INFORMATION:
APPLICANT: Ballinger, Dennis G.
APPLICANT: Ding, Wei
APPLICANT: Wagner, Susanne
APPLICANT: Hess, Mark A.
TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
FILE REFERENCE: Myriad 3
CURRENT APPLICATION NUMBER: US/09/262,773
CURRENT FILING DATE: 1999-03-04
NUMBER OF SEQ ID NOS: 210
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 210
LENGTH: 23071
TYPE: DNA
ORGANISM: human
US-09-262-773-210

Query Match
Best Local Similarity 100.0%; Score 25; DB 4; Length 23071;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2539 TGTCCAAAAA 2563
DB 920 TGTCCAAAAA 944

RESULT 11
US-09-345-882-1/C
Sequence 1, Application US/09345882
Patent No. 6399373
GENERAL INFORMATION:
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
FILE REFERENCE: GENSET.031A
CURRENT APPLICATION NUMBER: US/09/345,882
CURRENT FILING DATE: 1999-06-30
PRIOR APPLICATION NUMBER: US 60/091,315
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/111,909
PRIOR FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 140
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 162450
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 72794
OTHER INFORMATION: 5-124-273 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 88073 : polymorphic base A or C
OTHER INFORMATION: 5-127-261 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 90842
OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
FEATURE:
NAME/KEY: allele

LOCATION	93714	:		: polymorphic base deletion of GT
OTHER INFORMATION:	5-128-60	:		
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	97122	:		
OTHER INFORMATION:	99-1442-224	:		: polymorphic base G or T
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	97152	:		
OTHER INFORMATION:	5-129-144	:		: polymorphic base deletion of T
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	99098	:		
OTHER INFORMATION:	5-130-257	:		: polymorphic base A or G
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	99117	:		
OTHER INFORMATION:	5-130-276	:		: polymorphic base A or G
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	103806	:		
OTHER INFORMATION:	5-131-395	:		: polymorphic base A or T
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	106940	:		
OTHER INFORMATION:	5-133-375	:		: polymorphic base insertion of A
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	108106	:		
OTHER INFORMATION:	5-135-155	:		: polymorphic base insertion of A
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	108149	:		
OTHER INFORMATION:	5-135-198	:		: polymorphic base insertion of GTTT
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	108308	:		
OTHER INFORMATION:	5-135-357	:		: polymorphic base A or G
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	108471	:		
OTHER INFORMATION:	5-136-174	:		: polymorphic base C or T
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	134134	:		
OTHER INFORMATION:	5-140-120	:		: polymorphic base C or T
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	134362	:		
OTHER INFORMATION:	5-140-348	:		: polymorphic base insertion of A
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	134374	:		
OTHER INFORMATION:	5-140-361	:		: polymorphic base insertion of CA
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	146328	:		
OTHER INFORMATION:	5-143-84	:		: polymorphic base A or G
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	146345	:		
OTHER INFORMATION:	5-143-101	:		: polymorphic base A or C
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	150329	:		
OTHER INFORMATION:	5-145-24	:		: polymorphic base A or G
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	160031	:		
OTHER INFORMATION:	5-148-352	:		: polymorphic base G or T
FEATURE:		:		
NAME/KEY:	allele	:		
LOCATION:	72771..72817	:		

```

OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID22
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37

```

FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:
Query Match 1.0%; Score 25; DB 4; Length 162450;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2539 TGTCCAAAAA 2563
DB 161659 TGTCCAAAAA 161635
RESULT 12
US-08-776-944-9/c
Sequence 9, Application US/08776944
Patent No. 5925520
GENERAL INFORMATION:
APPLICANT: TULLEY, GILIAN
APPLICANT: SULLIVAN, KEVIN
TITLE OF INVENTION: NUCLEOTIDE SEQUENCING METHOD
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHIVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,944
FILING DATE: 14-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9417211.1
FILING DATE: 25-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: SCRAWFORD, ARTHUR R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 1498-92
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4006
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
ORGANELLE: Mitochondrion
US-08-776-944-9

Query Match 0.9%; Score 23; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2541 TCCAAAAA 2563
DB 32 TCCAAAAA 10

RESULT 13
US-07-920-281C-25
Sequence 25, Application US/07920281C
Patent No. 5739026
GENERAL INFORMATION:
APPLICANT: Garoff, Henrik
APPLICANT: Liljestrom, Peter
TITLE OF INVENTION: DNA Expression Systems Based on
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/920,281C
FILING DATE: 13-AUG-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 828-103P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-241-1300
TELEFAX: 703-241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Simliki Forest Virus
FEATURE:
NAME/KEY: /label=terminator
LOCATION: 1..80
OTHER INFORMATION: /note="3' terminal sequence of cDNA expression
OTHER INFORMATION: vector complementary to alphavirus genomic RNA"
US-07-920-281C-25

Query Match 0.9%; Score 23; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2541 TCCAAAAAAAAAAAAAAAAAAAA 2563
|||||
3 TCCAAAAAAAAAAAAAAAAAAAA 25

RESULT 14

US-08-466-277-25

Sequence 25, Application US/08466277

Patent No. 6180666

GENERAL INFORMATION:

APPLICANT: Garoff, Henrik

Liljestrom, Peter

TITLE OF INVENTION: DNA Expression Systems Based on

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,277

FILING DATE: 06-Jun-1995

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/920,281

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 828-103P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-241-1300

TELEFAX: 703-241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 80 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: RNA (genomic)

HYPOETHERICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Semliki Forest Virus

FEATURE:

NAME/KEY: -

LOCATION: 1..80

OTHER INFORMATION: /label= terminator

/note= "3' terminal sequence of CDNA expression

vector complementary to alphavirus genomic RNA"

SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-08-466-277-25

Query Match

Best Local Similarity 100.0%; Score 23; DB 4; Length 80;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2541 TCCAAAAAAAAAAAAAAAAAAAA 2563

|||||

3 TCCAAAAAAAAAAAAAAAAAAAA 25

RESULT 15

US-09-385-982-187/C

Sequence 187, Application US/09385982

Patent No. 6262334

GENERAL INFORMATION:

APPLICANT: ENDEGE, WILSON O., ET AL.

TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

FILE REFERENCE: CCNDA-260XX

CURRENT APPLICATION NUMBER: US/09/385,982

EARLIER FILING DATE: 1999-08-30

EARLIER APPLICATION NUMBER: 09/328,111

EARLIER FILING DATE: 1999-06-08

EARLIER APPLICATION NUMBER: 60/117,393

EARLIER FILING DATE: 1999-01-27

EARLIER APPLICATION NUMBER: 60/098,639

EARLIER FILING DATE: 1998-08-31

NUMBER OF SEQ ID NOS: 344

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 187

LENGTH: 256

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)...(256)

OTHER INFORMATION: n = A,T,C or G

US-09-385-982-187

Query Match

Best Local Similarity 100.0%; Score 23; DB 4; Length 256;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2541 TCCAAAAAAAAAAAAAAAAAAAA 2563

|||||

3 TCCAAAAAAAAAAAAAAAAAAAA 11

Search completed: April 20, 2003, 21:12:54

Job time: 185 secs

Db	121	CGCAGCTGACGCGAGGGGCGCGCAGAGGGGTGGCCATGCGTGTGGAGGGGCGCGGC	180
Qy	181	GGCGAGGGCGGGCGCCGACAGGGGGAAAGAGCGGGGGCGGGGCTCAGCCGCTGGCC	240
Db	181	GGCGAGGGCGGGCGCGCAGAGGGGGAAAGAGCGGGGGCGGGGCTCAGCCGCTGGCC	240
Qy	241	GGGGCGGGGGGAATGCTCATGCTGTACGCCATGCCGTGCCCGGGGTGGGGAGAGC	300
Db	241	GGGGCGGGGGGAATGCTCATGCTGTACGCCATGCCGTGCCCGGGGTGGGGAGAGC	300
Qy	301	TGAAGCAGGGCCAAAGAGATGAGAGAGCGGAGAGTACGTCCTTAAGTGGCCACCGACCA	360
Db	301	TGAAGCAGGGCCAAAGAGATGAGAGAGCGCGAGAGTACGTCCTTAAGTGGCCACCGACCA	360
Qy	361	AGGGCCCCAAGAACCAATCCAGTTGCTGATGATGACAGAGAGTCCACCAATTCGCCA	420
Db	361	AGGGCCCCAAGAACCAATCCAGTTGCTGATGATGACAGAGAGTCCACCAATTCGCCA	420
Qy	421	CCAAAATCGGCGCGAAGATCTTTGTGTGCTGCATCTCACAAGTCTCAGCTACAGCTACA	480
Db	421	CCAAAATCGGCGCGAAGATCTTTGTGTGCTGCAGTCTCACAAGTCTCAGCTACAGCTACA	480
Qy	481	GTCAGCTGCATCCTTACACAGATAGCGTCGATGATGATGATGATTTCTCCCGAGAGAGAGC	540
Db	481	GTCAGCTGCATCCTTACACAGATAGCGTCGATGATGATGATGATTTCTCCCGAGAGAGAGC	540
Qy	541	AAACCACTCCAAAGGCGAGCAGCAATTTCTGTGTAAGAAATCAAGCAGCGAGATTTG	600
Db	541	AAACCACTCCAAAGGCGAGCAGCAATTTCTGTGTAAGAAATCAAGCAGCGAGATTTG	600
Qy	601	GACGCCGGGAGATTGAGATTGACAGCAGACATGTCTCTCTGATTTCACTCAGAGAAC	660
Db	601	GACGCCGGGAGATTGAGATTGACAGCAGACATGTCTCTCTGATTTCACTCAGAGAAC	660
Qy	661	GTGCTCAGGGGGANAACCCCTTGGCTGGTGTAAATATGGGCTGTACACATCAAG	720
Db	661	GTGCTCAGGGGGANAACCCCTTGGCTGGTGTAAATATGGGCTGTACACATCAAG	720
Qy	721	CCCAAGACAGCGTGTATTGATTTAGACACATCTGTGCCCTGGGGCTCAGTCCCGCTGTG	780
Db	721	CCCAAGACAGCGGTATTGATTTAGACACATCTGTGCCCTGGGGCTCAGTCCCGCTGTG	780
Qy	781	CTTGTACATCTACTCACTCACTCAAGATGAAGTACTGACACACTGGCTGAGGCTGAGTTG	840
Db	781	CTTGTACATCTACTCACTCACTCAAGATGAAGTACTGACACACTGGCTGAGGCTGAGTTG	840
Qy	841	CAGTGTTCGCTTGGAAAGGGGAGTCAAGAGTACTTCTGTGTGTATTGACCGCTGTG	900
Db	841	CAGTGTTCGCTTGGAAAGGGGAGTCAAGAGTACTTCTGTGTGTATTGACCGCTGTG	900
Qy	901	TGAACATGAGATGGGTGGCAGCGCCAACTGATGCTGTGATGAGGGGAGACTTAACCCACT	960
Db	901	TGAACATGAGATGGGTGGCAGCGCCAACTGATGCTGTGATGAGGGGAGACTTAACCCACT	960
Qy	961	GGGTTTATAGAAAGTATCAAAAGTGTTTAAGAAATCCGAGGCAATTGTGGAAGAGACG	1020
Db	961	GGGTTTATAGAAAGTATCAAAAGTGTTTAAGAAATCCGAGGCAATTGTGGAAGAGACG	1020
Qy	1021	TGACTGCTTCCACAGGCTGTATCAGCTCTCCAAAGCTGGAGAGCTGTGTTCGGGCCA	1080
Db	1021	TGACTGCTTCCACAGGCTGTATCAGCTCTCCAAAGCTGGAGAGCTGTGTGTTCGGGCCA	1080
Qy	1081	TGAACGTCATGATTCGTTCATCCAAACAGAGTTTATATACCTTGTACTGTGCGAGAT	1140
Db	1081	TGAACGTCATGATTCGTTCATCCAAACAGAGTTTATATACCTTGTACTGTGCGAGAT	1140
Qy	1141	CCATTTGATGAGGCTCAAGAGCAGACAGATGTGATGTGGTGGGAAACAAGTGGTG	1200
Db	1141	CCATTTGATGAGGCTCAAGAGCAGACAGATGTGATGTGGTGGGAAACAAGTGGTG	1200
Qy	1201	TGTGTGCTATGTGAGTGTGGCAGAGGCTGTGTCTCTCTCAAAAGCTCTTGGAGCA	1260

Db	1201	TGTGTGGCTAATGGTGAAGTGAAGCAAGGGGCTGTGTGCTGTCTCTCAAAAGCTCTTGGAGCAA	1260
Qy	1261	TTGTCTACATTTACGGAATTCGACCCCATCTGTGCTCTGCAAGGCTTCGATCGATGGGTTCA	1320
Db	1261	TTGTCTACATTTACGGAATTCGACCCCATCTGTGCTCTGCAAGGCTTCGATCGATGGGTTCA	1320
Qy	1321	GGGTGTGAACCTTAATGAAAGTCAATCCGGCAAGTCGATGTCTGAATTAAGTTCTGCACAGGAA	1380
Db	1321	GGGTGTGAACCTTAATGAAAGTCAATCCGGCAAGTCGATGTCTGAATTAAGTTCTGCACAGGAA	1380
Qy	1381	ATTAAGATGTGTGACACGGGAGCACTGGATGCAATGAAAAACAATTTGATTCGATATGCA	1440
Db	1381	ATTAAGATGTGTGACACGGGAGCACTGGATGCAATGAAAAACAATTTGATTCGATATGCA	1440
Qy	1441	ATATGGGCACTCCACACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGAGAGT	1500
Db	1441	ATATGGGCACTCCACACACAGAAATCGATGTGACACAGCCTCCGCACTCCGAGCTGAGAGT	1500
Qy	1501	GGGACCGAGTACGTTTTCAGGTTGGACCATGTGCATCTGGCCAGATGGCCAAACAGATTTGTC	1560
Db	1501	GGGACCGAGTACGTTTTCAGGTTGGACCATGTGCATCTGGCCAGATGGCCAAACAGATTTGTC	1560
Qy	1561	TCTCTGGCAGAGGGTCGCTACTCAATTTGAGCGCTCCACAGTTCGCCACCTTGTGTTCTGT	1620
Db	1561	TCTCTGGCAGAGGGTCGCTACTCAATTTGAGCGCTCCACAGTTCGCCACCTTGTGTTCTGT	1620
Qy	1621	CCATCACAGCCACAACACAGGCTTTGGCAGCTGATAGAACTCTAATATGCAACCCGAGGGC	1680
Db	1621	CCATCACAGCCACAACACAGGCTTTGGCAGCTGATAGAACTCTAATATGCAACCCGAGGGC	1680
Qy	1681	GATACAGCAGAGATGTGTACTTGTCTTCTAAGAAATGATGATGATACGTTGCCAGCTTGC	1740
Db	1681	GATACAGCAGAGATGTGTACTTGTCTTCTAAGAAATGATGATGATACGTTGCCAGCTTGC	1740
Qy	1741	ATCTCCCATCATTTTGTGAGCCCACTTACAGAGCTGACAGATGACCAAGAAATATCTGG	1800
Db	1741	ATCTCCCATCATTTTGTGAGCCCACTTACAGAGCTGACAGATGACCAAGAAATATCTGG	1800
Qy	1801	GACTCAACAAAAATGGGCACTTCAAACTTAATTATACAGATACTAATGGACCATATAC	1860
Db	1801	GACTCAACAAAAATGGGCACTTCAAACTTAATTATACAGATACTAATGGACCATATAC	1860
Qy	1861	CAAGACACAGTCCACCTGAAACCAACACACTCTAAAGAAATATTTTTTAAGATTAAGTTTAT	1920
Db	1861	CAAGACACAGTCCACCTGAAACCAACACACTCTAAAGAAATATTTTTTAAGATTAAGTTTAT	1920
Qy	1921	TTTTCTCTACTCCTTTCCTTCTGTGATTTTTTTTCCCTAATATTCATCTCTGTTTTTCATC	1980
Db	1921	TTTTCTCTACTCCTTTCCTTCTGTGATTTTTTTTCCCTAATATTCATCTCTGTTTTTCATC	1980
Qy	1981	TCATATATCAAGTTCTGTGAGACCAACAGAGACCTTGCTTCAATGAGCTCTTTAATGAAATTA	2040
Db	1981	TCATATATCAAGTTCTGTGAGACCAACAGAGACCTTGCTTCAATGAGCTCTTTAATGAAATTA	2040
Qy	2041	GAAATTCAGGGTCCCTCACTCTAAGTCACTAAGAGAGATTTTACTCCCCACCCAGCCAGAA	2100
Db	2041	GAAATTCAGGGTCCCTCACTCTAAGTCACTAAGAGAGATTTTACTCCCCACCCAGCCAGAA	2100
Qy	2101	GGTGATCTTCTCTTTACCATTTCTGGGGCACTTACTCTTAATTAAGTACCTTAATTAAACA	2160
Db	2101	GGTGATCTTCTCTTTACCATTTCTGGGGCACTTACTCTTAATTAAGTACCTTAATTAAACA	2160
Qy	2161	GGAATTCGAAGGTAAGTCTCTCTGTGGAAACAAATCTGCAATGTCTAAATGCGCTTAAAGA	2220
Db	2161	GGAATTCGAAGGTAAGTCTCTCTGTGGAAACAAATCTGCAATGTCTAAATGCGCTTAAAGA	2220
Qy	2221	GGCCATTTCTTAGCTGCTGAAATACAGTGTCTTTTCACTTCTTCAGAGAAACAGGGATGTGT	2280
Db	2221	GGCCATTTCTTAGCTGCTGAAATACAGTGTCTTTTCACTTCTTCAGAGAAACAGGGATGTGT	2280
Qy	2281	ACCTTACCCGCGAGGTAGATTAGATGTGGGTGTGATGTTAAATTTCCCTTAGAATTTCCA	2340
Db	2281	ACCTTACCCGCGAGGTAGATTAGATGTGGGTGTGATGTTAAATTTCCCTTAGAATTTCCA	2340

Query Match	22.2%	Score 570	DB 10	Length 721
Best Local Similarity	99.8%	Pred. No. 1.5e-292		
Matches 620	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	53	CGCTGCGTGTCCGGCTGCTTGGCTGCGCAACACAAAGCGGTGGGCGCACAGCACTCA	112	
Db	56	CCGTCGCTGTCCGGCTGCTTGGCTGCGCAACACAAAGCGGTGGGCGCACAGCACTCA	115	
QY	113	GAAGCCGACGACGCTCGACGCGAGGGGCCGCGAGAGGCTGCGCTGTCCGAGGG	172	
Db	116	GAAGCCGACGACGCTCGACGCGAGGGGCCGCGAGAGGCTGCGCTGTCCGAGGG	175	
QY	173	CGCCGCGCGGGCGAGGGCGGGCGGCCCGACAGAGGGGGAAGAGCGCGGGCGCGCGGTCAAC	232	
Db	176	CGCCGCGCGGGCGAGGGCGGGCGGCCCGACAGAGGGGGAAGAGCGCGGGCGCGCGGTCAAC	235	
QY	233	CGCGCGCGCGGGCGCGCGGGGAATTTCATGCGCTACAGCGATGCGCGTCCCGGGGTCGG	292	
Db	236	CGCTGGCGGGCGGGCGGGGGAATTTCATGCTTACGCGATGCCGTCGCCGGGGTCCGG	295	
QY	293	GGAGGAGCTGAAGCAGGCGCAAGAGATCGAGGACGCCGAGAAGTACTCCTTCATGGCCAC	352	
Db	296	GGAGGAGCTGAAGCAGGCGCAAGAGATCGAGGACGCCGAGAAGTACTCCTTCATGGCCAC	355	
QY	353	CGTCACCAAGCGCGCCCAAGAAAGCAATCATCGTTTCTGTATGACATGCAAGATTCCACCA	412	
Db	356	CGTCACCAAGCGCGCCCAAGAAAGCAATCATCGTTTCTGTATGACATGCAAGATTCCACCA	415	

	Query Match	14.4%	Score 370;	DB 10;	Length 636;	
	Best Local Similarity	100.0%	Pred. No. 3.6e-186;			
	Matches 370;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	756 CTGGGGCGCTCAGTGCCTGGCTGCTCTTCTTAACACATCTACTCAACTCAGAAATGATAGCT	815				
Db	370 CTGGGGCGCTCAGTGCCTGGCTGCTCTTCTTAAACATCTACTCAACTCAGAAATGATAGCT	311				
QY	816 GCAGACATCGGCTGAGGCGCTGAGATTGCAGTGTTCGCTTGGAAAGGCGGAGTCAAGAAATGAC	875				
Db	310 GCAGACATCGGCTGAGGCGCTGAGATTGCATGTTCCTTGGAAAGGCGGAGTCAAGAAATGAC	251				
QY	876 TTCTGGTGTGTATTGACCCGCTGTGTGAACATGATGAGGTGGCAGGCCAACATGATCTGTG	935				
Db	250 TTCTGGTGTGTATTGACCCGCTGTGTGAACATGATGAGGTGGCAGGCCAACATGATCTGTG	191				


```
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16 2C1
CURRENT APPLICATION NUMBER: US/09/783,590
CURRENT FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
PRIOR FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 278
LENGTH: 140
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (10)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (71)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (117)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (125)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-278

Query Match
Best Local Similarity 1.0%; Score 45; DB 10; Length 140;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2519 CCTGAGATTATTTCTGTGTCACAAAAA 2563
DB 72 CCTGAGATTATTTCTGTGTCACAAAAA 116

RESULT 10
US-10-114-893-33
Sequence 33, Application US/10114893
Publication No. US20020193567A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Collins-Racle, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Bowman, Michael R.
APPLICANT: Spaulding, Vikki
APPLICANT: Carlin-Duckett, McKeough
APPLICANT: Kelleher, Kerry S.
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: GI 6000-10A
CURRENT APPLICATION NUMBER: US/10/114,893
CURRENT FILING DATE: 2002-04-02
EARLIER APPLICATION NUMBER: 09/413,232
EARLIER FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 321
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 33
LENGTH: 863
TYPE: DNA
ORGANISM: Homo sapiens
US-10-114-893-33

Query Match
Best Local Similarity 1.0%; Score 26; DB 9; Length 863;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2538 TTGTCCAAAAA 2563
DB 797 TTGTCCAAAAA 822

RESULT 11
US-09-770-445-261/C
Sequence 261, Application US/09770445
Patent No. US20020023281A1
GENERAL INFORMATION:
APPLICANT: Gorlach, Jorn
APPLICANT: An, Yong-Oiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krieker, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurlan, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2023US (PAR-012PRV)
CURRENT APPLICATION NUMBER: US/09/770,445
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: US 60/178,472
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 261
LENGTH: 987
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(987)
OTHER INFORMATION: n = A,T,C or G
US-09-770-445-261

Query Match
Best Local Similarity 1.0%; Score 26; DB 10; Length 987;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2538 TTGTCCAAAAA 2563
DB 38 TTGTCCAAAAA 13

RESULT 12
US-09-847-519A-1
Sequence 1, Application US/09847519A
Patent No. US20020102693A1
GENERAL INFORMATION:
APPLICANT: Lucie, Ralf M.
APPLICANT: Wei, Bo
TITLE OF INVENTION: DSP-14 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125,422
CURRENT APPLICATION NUMBER: US/09/847,519A
CURRENT FILING DATE: 2001-05-01
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1165
```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-847-519A-1

Query Match
Best Local Similarity 1.0%; Score 26; DB 10; Length 1165;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2538 TTGTCCAAAAAAAAAAAAAAAAAAAA 2563
Db 1139 TTGTCCAAAAAAAAAAAAAAAAAAAA 1164

RESULT 13
US-09-938-842A-3304
; Sequence 3304, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3304
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-3304

Query Match
Best Local Similarity 1.0%; Score 26; DB 9; Length 2000;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2538 TTGTCCAAAAAAAAAAAAAAAAAAAA 2563
Db 940 TTGTCCAAAAAAAAAAAAAAAAAAAA 965

RESULT 14
US-09-887-576-15/c
; Sequence 15, Application US/09887576
; Patent No. US2002014047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FASTSEQ for Windows version 4.0

```

```

; SEQ ID NO 15
; LENGTH: 2003
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-15

Query Match
Best Local Similarity 1.0%; Score 26; DB 10; Length 2003;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2538 TTGTCCAAAAAAAAAAAAAAAAAAAA 2563
Db 1009 TTGTCCAAAAAAAAAAAAAAAAAAAA 984

RESULT 15
US-09-783-590-4696
; Sequence 4696, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.201
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4696
; LENGTH: 101
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (89)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (91)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-4696

Query Match
Best Local Similarity 1.0%; Score 25; DB 10; Length 101;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2539 TTGTCCAAAAAAAAAAAAAAAAAAAA 2563
Db 60 TTGTCCAAAAAAAAAAAAAAAAAAAA 84

Search completed: April 20, 2003, 23:52:52
Job time : 292 secs

```

1

BASE COUNT	241 a	208 c	275 g	212 t	2 others
ORIGIN					

Oy	289	TCGGGGAGAGACTCTACAGAGCCCAAGAGATTCGAGAGCGCCGAGAGATCTCCTTAATGG	348
Db	1	TCGGGGAGAGACTCTACAGAGCCCAAGAGATTCGAGAGCGCCGAGAGATCTCCTTAATGG	60
Oy	349	CCACCGTCACCAAGCGGCCCAAGAACCAATTCAGTTTGCTGATGACATCAGAGATTCA	408
Db	61	CCACCGTCACCAAGCGGCCCCCAAGAACCAATTCAGTTTGCTGATGACATCAGAGATTCA	120
Oy	409	CCAAATTCGCCACCAAAACTGGCGGGAAGATCTTGTCGCTGATCTACAGTCCCTCCA	468
Db	121	CCAAATTCGCCACCAAAACTGGCGGGAAGATCTTGTCGCTGATCTACAGTCCCTCCA	180
Oy	469	CTGACAGCTACAGTTCAAGCTGCATCTACACAGATAGCTCTGATGATGAGGTTTCTCC	528
Db	181	CTGACAGCTACAGTTCAAGCTGCATCTACACAGATAGCTCTGATGATGAGGTTTCTCC	240
Oy	529	GAGAGAAACACCAAAACCACTCCCAAGGGCAGACAGCAATTTCTGTGTAAACAATCAAG	588
Db	241	GAGAGAAACACCAAAACCACTCCCAAGGGCAGACCAAAATTTCTGTGTAAACAATCAAG	300
Oy	589	AGGCAGATTTGGAGCGCCGGAGATTAGATTGCAGACCAAGACATGTCTGCTGATTT	648
Db	301	AGGCAGATTTGGAGCGCCGGGAGATTAGATTGCAGACCAAGACATGTCTGCTGATTT	360
Oy	649	CACATCAGAAACGTGCTCAGGGGAGAAAGCCCTTGCTGGTGTCTAAATATGTGGGCTGTA	708
Db	361	CACATCAGAAACGTGCTCAGGGGAGAAAGCCCTTGCTGGTGTCTAAATATGTGGGCTGTA	420
Oy	709	CACATCAGAGCCACAGACAGCGGTGTGATTTGAGACACTGTGCGCCCTGGGGGCTCAGT	768
Db	421	CACATCAGAGCCACAGACAGCGGTGTGATTTGAGACACTGTGCGCCCTGGGGGCTCAGT	480

QY	769	GCCTGGTGTGCTTGAATCACTACCACTAGAGTGAAGTACGACGACTGGCTG	828
Db	481	GCCTGGTGTGCTTGTWACCTTACTCACTCAAGTGAAGTACGACGACTGGCTG	540
QY	829	AGCGTGAAGTTGACAGTGTTCCTTGGAAAGGCGGTGCAAGATGACTTGTGGTGTGA	888
Db	541	AGCGTGAAGTTGACAGTGTTCCTTGGAAAGGCGAGTCAAGATGACTTGTGGTGTGA	600
QY	889	TTGACCGCTGTGTSAACATGGATGGGTGGCAGGCCAACATGATCTGGATGATGGGGAG	948
Db	601	TTGACCGGTGTGTSAACATGGATGGGTGGCAGGCCAACATGATCTGGATGATGGGGAG	660
QY	949	ACTTAAACCCACTGGGTTTATTAAGAGATATCCAAACGCTGTTTAAAGAGATCCGAGGACATTTG	1000
Db	661	ACTTAAACCCACTGGGTTTATTAAGAGATATCCAAACGCTGTTTAAAGAGATCCGAGGACATTTG	720
QY	1009	TGGAAGAGAGGTACATCGGTGTTCAACAGGCTGATCAAGCTCTCCAAAGCTGGGAAGCTCT	106
Db	721	TGGAAGAGAGGTACATCGGTGTTCAACAGGCTGATCAAGCTCTCCAAAGCTGGGAAGCTCT	780
QY	1069	GTGTTCGGGCATGAAGGTAAATATCTTTACCAACAGAAAGTTTGAAATTTGTACT	112
Db	781	GTGTTCGGGCATGAAGGTAAATATCTTTACCAACAGAAAGTTTGAAATTTGTACT	840

RESULT 2	LOCUS	DEFINITION
AL551097	AL551097	921 bp mRNA
AL551097	L1.NFL006.PL2	1linear EST 16-FEB-2001
AL551097	L1.NFL006.PL2	clone CS0100661M07 5
AL551097	L1.NFL006.PL2	prime, mRNA sequence.

Mamma

COMMENT	Contact	Genos

Email

source

BASE COUNT . 2
ORIGIN

Query Match

Matches 887;

27	00000000
26	11111111
25	00000000

QY 129 GACGCAG

Db 61 GACGCAG

0Y 689 GGCGCGGG
|||||

OV 249 GGGGGGAA

Db . 181 CGGGGAA

QY 309 GCCAAGG |.|||||

1
2
3
4
5
6
7
8

```

QY 369 AAGAGCAATCCAGTTGCTGATGACATGAGAGATTCCCAATTTCCCAACCAAACT 428
DB 301 AAGAGCAATCCAGTTGCTGATGACATGAGAGATTCCCAATTTCCCAACCAAACT 360
QY 429 GGGGAGAGATCTTGTGCTGCTGATCTCAGTCCCTCACTGACAGTACAGTTCACT 488
DB 361 GGGGAGAGATCTTGTGCTGCTGATCTCAGTCCCTCACTGACAGTACAGTTCACT 420
QY 489 GCATCTACACAGATAGCTGTGATGATGAGTTTCTCCCGAGAGAAACAGCAACAC 548
DB 421 GCATCTACACAGATAGCTGTGATGATGAGTTTCTCCCGAGAGAAACAGCAACAC 480
QY 549 TCCAGAGGAGAGCAAAATTTCTGTGTGAAGAACATCAAGAGAGAAATTTGAGCCGG 608
DB 481 TCCAGAGGAGAGCAAAATTTCTGTGTGAAGAACATCAAGAGAGAAATTTGAGCCGG 540
QY 609 GAGATTGAGATTGAGAGCAAGATGCTGCTGTGATTTTACATCAGAGAACTGCTAG 668
DB 541 GAGATTGAGATTGAGAGCAAGATGCTGCTGTGATTTTACATCAGAGAACTGCTAG 600
QY 669 GGGGAGAGAGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 728
DB 601 GGGGAGAGAGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 729 GCGGTGTGATTGAGACACTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 788
DB 661 GCGGTGTGATTGAGACACTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 789 ATCTACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 848
DB 721 ATCTACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 849 GCTTGAAGAGGCGAGTGCAGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 908
DB 781 GCTTGAAGAGGCGAGTGCAGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 909 GATGGGTGCGAGGCGCAACATGATCTGATGATGATGATGATGATGATGATGATGAT 956
DB 841 GATGGGTGCGAGGCGCAACATGATCTGATGATGATGATGATGATGATGATGATGAT 888

```

```

RESULT 3
LOCUS U69200 1379 bp mRNA linear EST 27-OCT-1999
DEFINITION U69200 Soares infant brain INIB Homo sapiens cDNA clone 23931, mRNA
sequence.
ACCESSION U69200
VERSION U69200.1 GI:2739423
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1379)
Volorio, S., Simon, G., Repetto, M., Cucetardi, M., Banfi, S., Borani,
G., Ballabio, A. and Zollo, M.
Sequencing analysis of forty-eight human image cDNA clones similar
to desophila mutant protein
JOURNAL DNA Seq. 9 (5-6), 307-315 (1998)
MEDLINE 99452368
COMMENT CONTACT: Zollo, Massimo
Flethron Institute of Genetics and Medicine
Via Olgettina 58, Milan, MI 20132, Italy
Email: zollo@tigem.it

```

FEATURES

```

source
1..1379
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1p13.3"
/clone="23931"
/clone.lib="Soares infant brain INIB"
/sex="female"
/dev_stage="73 days post. natal"

```

```

/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: latmid B; Site: 1: Not
I; Site: 2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5',
AAGTGAAGATTTGCGGCGCCAGAGATTTTTTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the latmid B vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 372 a 315 c 292 g 400 t
ORIGIN

```

```

Query Match 32.2%; Score 826; DB 14; Length 1379;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 826; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1201 TGTGTGCTATGATGAGTAGAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
DB 3 TGTGTGCTATGATGAGTAGAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 62
QY 1261 TTGTCTATTTACCAATCCAGCCCATCTGTGCTCTGAGGCTGCTGATGATGGTTCA 1320
DB 63 TTGTCTATTTACCAATCCAGCCCATCTGTGCTCTGAGGCTGCTGATGATGGTTCA 122
QY 1321 GGGTGTAAAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 123 GGGTGTAAAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 182
QY 1381 ATTAAGATGTAGTGAACAGGAGACACTTGATGATGATGATGATGATGATGATGATGAT 1440
DB 183 ATTAAGATGTAGTGAACAGGAGACACTTGATGATGATGATGATGATGATGATGATGAT 242
QY 1441 ATATGGGCGATCCCAACACAGAAATGATGATGATGATGATGATGATGATGATGAT 1500
DB 243 ATATGGGCGATCCCAACACAGAAATGATGATGATGATGATGATGATGATGATGAT 302
QY 1501 GGGAGCGAGTACGCTGCTGAGTGAACATGATGATGATGATGATGATGATGATGATGAT 1560
DB 303 GGGAGCGAGTACGCTGCTGAGTGAACATGATGATGATGATGATGATGATGATGATGAT 362
QY 1561 TCTGTGCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
DB 363 TCTGTGCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
QY 1621 CCATACAGCCACACACAGAGTTTGGCACTGATGATGATGATGATGATGATGATGATGAT 1680
DB 423 CCATACAGCCACACACAGAGTTTGGCACTGATGATGATGATGATGATGATGATGATGAT 482
QY 1681 GATACAGACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
DB 483 GATACAGACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
QY 1741 ATCTGCCATCTTGTGATCCCACTTACAGAGTGCAGAGTGCAGAGTGCAGAGTGCAG 1800
DB 543 ATCTGCCATCTTGTGATCCCACTTACAGAGTGCAGAGTGCAGAGTGCAGAGTGCAG 602
QY 1801 GATCTCAAAAAAAGGCGCATTCCTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860
DB 603 GATCTCAAAAAAAGGCGCATTCCTAATTAATTAATTAATTAATTAATTAATTAATTAAT 662
QY 1861 CAAGGACAGTGCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
DB 663 CAAGGACAGTGCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 722
QY 1921 TTCTCTTACTCTTCTCTCTGATTTTCTGATTAATTAATTAATTAATTAATTAATTAAT 1980
DB 723 TTCTCTTACTCTTCTCTCTGATTTTCTGATTAATTAATTAATTAATTAATTAATTAAT 782
QY 1981 TCATTATCAAGTCTGAGACACACAGAACTGCTTCATGGCT 2026
DB 783 TCATTATCAAGTCTGAGACACACAGAACTGCTTCATGGCT 828

```

```

RESULT 4
BM453920      867 bp      mRNA      linear      EST 05-FEB-2002
LOCUS         AGENCOURT.6402958 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528955
DEFINITION   5', mRNA sequence.
ACCESSION    BM453920
VERSION      BM453920.1 GI:18502949
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE    1 (bases 1 to 867)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgarbs-remail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTF
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LLM412206 row: 1 column: 20
            High quality sequence stop: 670.
            Location/Qualifiers
                1..867
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_1db="NIH_MGC_72"
                /tissue_type="melanotic melanoma"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally. Primer: oligo dt.
                Average insert size 2 kb. Library constructed by Life
                Technologies."
BASE COUNT   222 a      201 c      241 g      202 t      1 others
ORIGIN
Query Match      30.7%; Score 788; DB 13; Length 867;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 328 CCGAGAGTACTCTTATGCGCCGCTCCACCAAGGCGCCCAAGCAATCCAGTTTG 387
      |||||||
DB 1 CCGAGAGTACTCTTATGCGCCGCTCCACCAAGGCGCCCAAGCAATCCAGTTTG 60

OY 388 CTGATGACATGACAGAGTTCACCAATTCGCCACCAAAATGCGGAAGATCTTGCTC 447
      |||||||
DB 61 CTGATGACATGACAGAGTTCACCAATTCGCCACCAAAATGCGGAAGATCTTGCTC 120

OY 448 GCTCGATCTCACAGTCTCCACTGACAGCTACAGTTCAGTCTGATCTACAGATAGCT 507
      |||||||
DB 121 GCTCGATCTCACAGTCTCCACTGACAGCTACAGTTCAGTCTGATCTACAGATAGCT 180

OY 508 CTGATGATGAGTTTCTCCCGGAGAGCAAACTCCAAAGGCGCGCAAT 567
      |||||||
DB 181 CTGATGATGAGTTTCTCCCGGAGAGCAAACTCCAAAGGCGCGCAAT 240

OY 568 TCTGTGAGAGACATCATCAGAGGACAAATTTGGAGCCGGAGATTCAGATTCAGAGC 627
      |||||||
DB 241 TCTGTGAGAGACATCATCAGAGGACAAATTTGGAGCCGGAGATTCAGATTCAGAGC 300

OY 628 AAGACATGCTGCTCTGATTTACTCAGAGAAAGCTGCTCAGGGGAGAAAGCCCTTGCTG 687
      |||||||
DB 301 AAGACATGCTGCTCTGATTTACTCAGAGAAAGCTGCTCAGGGGAGAAAGCCCTTGCTG 360

OY 688 GTGCTAAATAGTGGCTGTACACATCAGCCGAGAGCGGCTGTGATAGACAC 747
      |||||||

```

```

DB 361 GTGCTAAATAGTGGCTGTACACATCAGCCAGAGCGGCTGTGATAGACAC 420
OY 748 TCTGTGCCCCCTGGGGGCTCAGTGGCGCTGTGCTTTATCATCTACTCACTAGAAATG 807
DB 421 TCTGTGCCCCCTGGGGGCTCAGTGGCGCTGTGCTTTATCATCTACTCACTAGAAATG 480
OY 808 AAGTACGTGACAGCACTGCTGAGGCTGAGATTGAGATGTTGCTTGGAAGGCGAGTCAG 867
DB 481 AAGTACGTGACAGCACTGCTGAGGCTGAGATTGAGATGTTGCTTGGAAGGCGAGTCAG 540
OY 868 AAGATGACTTCTGTGTTGATTTAGACCGCTGTGTAACATGATGGTGGCAGGCCAACA 927
DB 541 AAGATGACTTCTGTGTTGATTTAGACCGCTGTGTAACATGATGGTGGCAGGCCAACA 600
OY 928 TGATCTGTGATGATGGGGGAGACATTAAACCCACTGGGTTTAAAGATATCCAAAGCT 987
DB 601 TGATCTGTGATGATGGGGGAGACATTAAACCCACTGGGTTTAAAGATATCCAAAGCT 660
OY 988 TTAGAAGATCCGAGGCGCTGTGGAAGAGACGCTGACTGCTTCACAGGCTGTATCAGC 1047
DB 661 TTAGAAGATCCGAGGCGCTGTGGAAGAGACGCTGACTGCTTCACAGGCTGTATCAGC 720
OY 1048 TCTCCAAAGCTGGGAGAGCTCTGTTCGCGCCATGAGTCAATGATCTGTACCAAAAC 1107
DB 721 TCTCCAAAGCTGGGAGAGCTCTGTTCGCGCCATGAGTCAATGATCTGTACCAAAAC 780
OY 1108 AGAAGTTT 1115
DB 781 AGAAGTTT 788

RESULT 5
BO212518      839 bp      mRNA      linear      EST 02-MAY-2002
LOCUS         AGENCOURT.7595081 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6062515
DEFINITION   5', mRNA sequence.
ACCESSION    BO212518
VERSION      BO212518.1 GI:20392818
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE    1 (bases 1 to 839)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgarbs-remail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTF
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LLM413334 row: b column: 20
            High quality sequence stop: 671.
            Location/Qualifiers
                1..839
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_1db="NIH_MGC_72"
                /tissue_type="melanotic melanoma"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally. Primer: oligo dt.
                Average insert size 2 kb. Library constructed by Life
                Technologies."
BASE COUNT   228 a      203 c      178 g      229 t      1 others
ORIGIN

```



```

DB 726 CCGGGTTTAAAGATATCAAGCTGTTTAAAGATCCGAGGATGTGGAAGAG 785
QY 1019 CCGTACTGTTTACAGAGCTGTATAGCTCTCCAAAGCTGGAGACTGTGTCCGGC 1078
DB 786 CCGTACTGTTTACAGAGCTGTATAGCTCTCCAAAGCTGGAGACTGTGTCCGGC 845
QY 1079 CATTGAGCT 1087
DB 846 CATTGAGCT 854

RESULT 7
B0059373 1021 bp mRNA linear EST 29-MAR-2002
LOCUS AGENCOURT_7050082 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5815413
DEFINITION 5', mRNA sequence.
ACCESSION B0059373
VERSION B0059373.1 GI:19818713
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL Unpublished (1999)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rt@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM2070 row: 3 column: 22
High quality sequence stop: 688.
Location/Qualifiers
1. 1021
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5815413"
/clone_lib="NIH_MGC_99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lymph. Vector: pOTB7. Site: 1: xhoI; site: 2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAGC(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using zap-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 284 a 245 c 201 g 291 t
ORIGIN
Query Match 28.7%; Score 735; DB 14; Length 1021;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1260 ATTGCTACATTACCGAAATGCAAGCCATCTGTCTCGAGGCTGATGAGGGTTC 1319
DB 1 ATTGCTACATTACCGAAATGCAAGCCATCTGTCTCGAGGCTGATGAGGGTTC 60
QY 1320 AGGGTGTAAAGCTAAATGAAGCATCGGCAAGTGATGTATTAATCTGCAAGGA 1379
DB 61 AGGGTGTAAAGCTAAATGAAGCATCGGCAAGTGATGTATTAATCTGCAAGGA 120
QY 1380 AATAAGATGTAGTGCACAGGAGCACTGTGATCGCATGAAAAAAGCTGTATCTATGC 1439

```

```

DB 121 AATAAGATGTAGTGCACAGGAGCACTGTGATCGATGAAAAAAGCTGTATCTATGC 180
QY 1440 AATATGGCCATCTCCAAACACAGAAATGATGACCAAGCTCCGCACTCCGGAGCTAGC 1499
DB 181 AATATGGCCATCTCCAAACACAGAAATGATGACCAAGCTCCGCACTCCGGAGCTAGC 240
QY 1500 TGGAGAGGAGTACGTCTCAGGTGAGACATGTATCTGGCCAGATGGCAAGAGTGTG 1559
DB 241 TGGAGAGGAGTACGTCTCAGGTGAGACATGTATCTGGCCAGATGGCAAGAGTGTG 300
QY 1560 CTCCTGGCAGAGGCTGCTACTCAATTTAGCTGCTCAGACAGTTCACACTTGTGTCTG 1619
DB 301 CTCCTGGCAGAGGCTGCTACTCAATTTAGCTGCTCAGACAGTTCACACTTGTGTCTG 360
QY 1620 TCCATCAGACCCACACACAGAGCTTTGGCAGCTGTAGACTTATTAATGCAAGCCGAGGG 1679
DB 361 TCCATCAGACCCACACACAGAGCTTTGGCAGCTGTAGACTTATTAATGCAAGCCGAGGG 420
QY 1680 CGATACAGCAGAGATGTGTACTTGTCTTCTTAAAGAAATGATGATACCTTGGCAGCTTG 1739
DB 421 CGATACAGCAGAGATGTGTACTTGTCTTCTTAAAGAAATGATGATACCTTGGCAGCTTG 480
QY 1740 CATCTGCCATCATTTTGTATGCGCCACCTTACAGAGCTGACAGTGAAGCAAGCAAAATATCTG 1799
DB 481 CATCTGCCATCATTTTGTATGCGCCACCTTACAGAGCTGACAGTGAAGCAAGCAAAATATCTG 540
QY 1800 GGACTCAACAAAATGGGCGCTTCAAACTATTTATTACAGATTAATGAGACATCTA 1859
DB 541 GGACTCAACAAAATGGGCGCTTCAAACTATTTATTACAGATTAATGAGACATCTA 600
QY 1860 CCAAGGACGAGTCCAGCTGAGACACACACTTAAGAAATTTTAAAGTAACTTTA 1919
DB 601 CCAAGGACGAGTCCAGCTGAGACACACACTTAAGAAATTTTAAAGTAACTTTA 660
QY 1920 TTTTCTTCTTACTCTCTTCTCTCTGATTTTTCCTAATTTCAATCTTGTCTTTCAT 1979
DB 661 TTTTCTTCTTACTCTCTTCTCTCTGATTTTTCCTAATTTCAATTTCTTGTCTTTCAT 720
QY 1980 CTCATTATCCAGCT 1994
DB 721 CTCATTATCCAGCT 735

RESULT 8
B0449470 1042 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6400881 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5493677
DEFINITION 5', mRNA sequence.
ACCESSION B0449470
VERSION B0449470.1 GI:18498510
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL Unpublished (1999)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rt@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM1217 row: m column: 06
High quality sequence stop: 694.
Location/Qualifiers
1. 1042
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
 /clone="IMAGE:5493677"
 /clone.lib="NIH_MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-Sport6; Site: 1: Not; Site: 2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."

BASE COUNT 267 a 219 c 309 g 240 t 7 others

Query Match 28.1%; Score 719; DB 13; Length 1042;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 769; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

532 AGAAGCAGCAAAACCACTCCAGGCGACAGCAATTTCTGTGAGAGAACATCAAGCAG 591
 1 AGAAGCAGCAAAACCACTCCAGGCGACAGCAATTTCTGTGAGAGAACATCAAGCAG 60
 592 CAGAAATTTGAGCGCGGAGATTGAGATTGACAGAGCAACATGTCTGTCTGATTTCAC 651
 61 CAGAAATTTGAGCGCGGAGATTGAGATTGACAGAGCAACATGTCTGTCTGATTTCAC 120
 652 TCAGGAACGCTCTCAGGCGGAGAAACCCCTTGCTGCTAAATAGTGGCTGTACAC 711
 121 TCAGGAACGCTCTCAGGCGGAGAAACCCCTTGCTGCTAAATAGTGGCTGTACAC 180
 712 ACATCACACCCAGACAGCGGTGTATGACACACTCTGCCCCGTGGGCTCAGTCC 771
 181 ACATCACACCCAGACAGCGGTGTATGACACACTCTGCCCCGTGGGCTCAGTCC 240
 772 GCGGCTCTGTATACATCTACTCACTCAGATGAGATGAGTGCAGCACTGGCTGAG 831
 241 GCTGCTCTGTATACATCTACTCACTCAGATGAGATGAGTGCAGCACTGGCTGAG 300
 832 CTGGAATTCAGTGTTCGCTTGGAGAGCGAGTCAAGATGACTTCTGTGTGTATG 891
 301 CTGGAATTCAGTGTTCGCTTGGAGAGCGAGTCAAGATGACTTCTGTGTGTATG 360
 892 ACCGCTGTGTAACATGATGGTGGCAGCCCAACATGATCCGTGATGAGGGGAGACT 951
 361 ACCGCTGTGTAACATGATGGTGGCAGCCCAACATGATCCGTGATGAGGGGAGACT 420
 952 TAACCCAGTGGGTTTAAAGATATCCAAAGGTTTAAAGATCCGAGGCTGTGG 1011
 421 TAACCCAGTGGGTTTAAAGATATCCAAAGGTTTAAAGATCCGAGGCTGTGG 480
 1012 AAGAGAGCTGATGCTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGCTGTG 1071
 481 AAGAGAGCTGATGCTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGCTGTG 540
 1072 TTCGGGCAATGAGCTGATGATCTGTACCAAGAGAGTTGATTAAGTCTGACTGCT 1131
 541 TTCGGGCAATGAGCTGATGATCTGTACCAAGAGAGTTGATTAAGTCTGACTGCT 600
 1132 GCCGAATATCATTTTGGATGGCTGAGAGGACACAGATGTGATTTTGGTGGAAAC 1191
 601 GCCGAATATCATTTTGGATGGCTGAGAGGACACAGATGTGATTTTGGTGGAAAC 660
 1192 AAGTGTGCTGTGCTGTATGAGTGTAGAGTGAAGGCTGCTGTGCTCTCAAGCTC 1251
 661 AAGTGTGCTGTGCTGTATGAGTGTAGAGTGAAGGCTGCTGTGCTCTCAAGCTC 720
 1252 TTGAGCAATTTCTATCATTTACGAATTCAGCCCATCTGTCTCTCAG 1301
 721 TTGAGCAATTTCTATCATTTACGAATTCAGCCCATCTGTCTCTCAG 770

RESULT 9
 B0896777 894 bp mRNA linear EST 16-AUG-2002
 LOCUS B0896777
 DEFINITION AGENOURT_8715624 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6425821

5', mRNA sequence.
 B0896777
 B0896777.1 GI:22288791
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (Bases 1 to 894)
 NIH-MGC http://mgs.ncl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 AUTHORS Contact: Robert Strausberg, Ph.D.
 COMMENT Email: c9abs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LICM2608 row: d column: 14
 High quality sequence stop: 700.
 Location/Qualifiers
 1. 894
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6425821"
 /clone.lib="NIH_MGC_101"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pCMV7; Site: 1: EcoRI; Site: 2: XhoI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(s). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 226 a 205 c 247 g 216 t

Query Match 27.0%; Score 693; DB 14; Length 894;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 693; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

910 ATGGGTGCGAGCCCAACATGATCTCGATGATGGGAGAGACTTAAACCACTGGCTTATA 969
 1 ATGGGTGCGAGCCCAACATGATCTCGATGATGGGAGAGACTTAAACCACTGGCTTATA 60
 970 AGAAGTATCCAAAGCTGTTAAGAGATCCGAGGATTTGGAGAGAGCGTGAATG 1029
 61 AGAAGTATCCAAAGCTGTTAAGAGATCCGAGGATTTGGAGAGAGCGTGAATG 120
 1030 TTCAAGGCTGTATCAGCTCTCCAAAGCTGGGAAAGCTGTGTGTCGGCCATGAACGTCA 1089
 121 TTCAAGGCTGTATCAGCTCTCCAAAGCTGGGAAAGCTGTGTGTCGGCCATGAACGTCA 180
 1090 ATGATTCGTTACCAACAGAGATTGATTAACCTGTACTGCTGCCGAATTCATTGG 1149
 121 TTCAAGGCTGTATCAGCTCTCCAAAGCTGGGAAAGCTGTGTGTCGGCCATGAACGTCA 180
 1150 ATGGCTGTAAAGAGCCACATGATGATGATTTGGGGAACAATGAGTGGTGGCT 1209
 241 ATGGCTGTAAAGAGCCACATGATGATGATTTGGGGAACAATGAGTGGTGGCT 300
 1210 ATGGAGAGTGAAGGAGGCTGTGCTGTCTCAAGAGCTTTGGAGCAATTTGTCTACA 1269
 301 ATGGAGAGTGAAGGAGGCTGTGCTGTCTCAAGAGCTTTGGAGCAATTTGTCTACA 360
 1270 TTACCGAATTCAGACCCCATCTGTGCTGTGACAGGCTGCATGAGATGGGTTCAAGGTTGTA 1329
 361 TTACCGAATTCAGACCCCATCTGTGCTGTGACAGGCTGCATGAGATGGGTTCAAGGTTGTA 420

QY	1330	AGC AAAATGAAGTATCCGGGAAGTCATGTCGTAATACCTTGACACAGAAATAGAAATG	1389
Dd	421	AGCTAAATGAAGTATCCGGCAAGTCATGTCGTAATACCTTGACACAGAAATAGAAATG	480
QY	1390	TAGTCACACGGGAGCACTTGGATGCGATGAATAAACAGTTGTATCGTATGCAATATGGGCC	1449
Dd	481	TAGTCACACGGGAGCACTTGGATGCGATGAATAAACAGTTGTATCGTATGCAATATGGGCC	540
QY	1450	ACTCCACACAGAAATGCATTTGACACACCTCCGACCTCCGGAGGTGACGTGGAGCCAG	1509
Dd	541	ACTCCACACAGAAATGCATTTGACACACCTCCGACCTCCGGAGGTGACGTGGAGCCAG	600
QY	1510	TACGTTTCAGGTGACCATGTCATCTGGCCAGATGGCAAAAGATTTCTCTCCGCGAG	1569
Dd	601	TACGTTTCAGGTGACCATGTCATCTGGCCAGATGGCAAAAGATTTCTCTCCGCGAG	660
QY	1570	AGGCTGCTACTCAATTGAGCGCTCCACAG	1602
Dd	661	AGGCTGCTACTCAATTGAGCTCCACAG	693

RESULT	10
LOCUS	BM468564
DEFINITION	BM468564 967 bp mRNA linear EST 05-PEB-2002 AGNCNCOURT 6475681 NIH_MGC_72 Homo sapiens CDNA IMAGE:5578170
ACCESSION	BM468564
VERSION	BM468564
KEYWORDS	5' UTR, mRNA sequence.
SOURCE	EST. human.

ORGANISM	<i>Homo sapiens</i> Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 967)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Email: cgapds-remail.nih.gov
Tissue Procurement: ATCC/DCFD/DMF
CDNA Library Preparation: Jaffe Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
plate: LAM12332 row: m column: 19
High quality sequence stop: 685.
Location/Qualifiers
1. .967

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone IMAGE:5578170"
/clone_1id="NH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/site="Organ: skin; Vector: pCMV-SPORT6, Site_1: NotI,
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
```

QY 1 GGGCGGGGAGGTGAGACTGGAGTGTCTTCTGGTTCTCTTGGGCACCGTCGCT 60
 |||||
 Db 3 GGGCGGGGAGGTGAGACTGGAGTGTCTTCTGGTTCTCTTGGGCACCGTCGCT 62
 |||||
 QY 61 GTTCGGGTCTTGGGCTGCCGAACAGCAAGGGCTGGGCCACAGCACTTAGAAGCCGA 120
 |||||

Db	63	GTCCGGCTGCTTGGGCTGTGCCGAACAGACAAAGGGCTGGGCCACAGCACTCTAGAAAGCCGA	122
QY	121	CGCAGCTTGACGCGAGGGGCCCGGACAGAGGGGTGGCGATCGGTGTGCGAAGGGCGCCGGC	180
Db	123	CGCAGCTCGAGCGAGGGGCCCGGACAGAGGGGTGGCGATCGGTGTGCGAAGGGCGCCGGC	182
QY	181	GGGCGAGGGGGGGGGCGGCAGAGGGGGGAAAGAGGGGGGGGGGGGGGGGATCGCGCTTGGGC	240
Db	183	GGGCGAGGGGGGGGGGGCGGCAGAGGGGGGAAAGAGGGGGGGGGGGGGGGGATCGCGCTTGGGC	242
QY	241	GGGCGCGGGGGGGAATGTGATGCTTGACGCGCATGCCCTGCCCGGGGTGCGGGAGAGAC	300
Db	243	GGGCGCGGGGGGGAATGTGATGCTTGACGCGCATGCCCTGCCCGGGGTGCGGGAGAGAC	302
QY	301	TGAAGCAGGCCCAAGAGATCCAGAGACGCCGAGAACTACTCTTCAATGGCCACCGTACCA	360
Db	303	TGAAGCAGGCCCAAGAGATCCAGAGACGCCGAGAACTACTCTTCAATGGCCACCGTACCA	362
QY	361	AGGCGCCCAAGAGAAATCCAGTTTGGTGAATGATCAGAGATCAGAGATTCACCAATTCGCCA	420
Db	363	AGGCGCCCAAGAGAAATCCAGTTTGGTGAATGATCAGAGATCAGAGATTCACCAATTCGCCA	422
QY	421	CCAAAATGTCGGCGAAGATCTTTGTCTCGTCGATCTCACAGTCTCCACTGACACTTACA	480
Db	423	CCAAAATGTCGGCGAAGATCTTTGTCTCGTCGATCTCACAGTCTCCACTGACACTTACA	482
QY	481	GTTGAGCTGCATCTTCACACAGATTAAGTCTGATGATGAGGTTCTCCCGAGAGAGCAC	540
Db	483	GTTGAGCTGCATCTTCACACAGATTAAGTCTGATGATGAGGTTCTCCCGAGAGAGCAC	542
QY	541	AAACCACTCCAAAGGGCAGACAGCAATTTCTGTGTAGAAACATCAAGCAGCAGAAATTTG	600
Db	543	AAACCACTCCAAAGGGCAGACAGCAATTTCTGTGTAGAAACATCAAGCAGCAGAAATTTG	602
QY	601	GACGCCGGGAGATTAGATTGCAGACAGCAAGACATGTCTGCTGTGATTTCACTCAGGAAC	660
Db	603	GACGCCGGGAGATTAGATTGCAGACAGCAAGACATGTCTGCTGTGATTTCACTCAGGAAC	662
QY	661	GGTCTCAGGGGAGAGAGCCCTTGCTGGTGGCTTGAATATGTTGGGTGTACACACATCACAG	720
Db	663	GGTCTCAGGGGAGAGAGCCCTTGCTGGTGGCTTGAATATGTTGGGTGTACACACATCACAG	722
QY	721	CCCAGACAGCGGTGTTGATTGAGACACTGTGCCCCGTGGGGGCTCAGTGCCTGGTCTG	780
Db	723	CCCAGACAGCGGTGTTGATTGAGACACTGTGCCCCGTGGGGGCTCAGTGCCTGGTCTG	782
QY	781	CTTG 784	
Db	783	CTTG 786	

RESULT 11	
A0131962	
LOCUS	
DEFINITION	A0131962 745 bp
ACCESSION	A0131962 NT2RP3 Homo sapiens cDNA
VERSION	A0131962
KEYWORDS	sequence.
SOURCE	A0131962.1 GI:10992316
ORGANISM	EST.
	human.
	Homo sapiens
	clone NT2RP3003545 5', mRNA
	1linear EST 01-AUG-2002

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eutheria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 745)
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Satou, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S., and Iwagaki, T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao, Isogat
Genomics Laboratory
Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@nri.co.jp

HRI human cDNA project; 5' & 3' end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES

source

Location/Qualifiers

1. 745

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="NT2RP1003545"

/clone_1db="NT2RP3"

/cell_type="teratocarcinoma"

/cell_line="NT2"

/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"

BASE COUNT 193 a 149 c 214 g 186 t 3 others

ORIGIN

Query Match 26.4%; Score 676; DB 9; Length 745;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 777 TCTGCTTGAACATCTACCACTCACTGAGTGAAGTGCAGCAGCTGCTGAGGCTGGA 836
DB 1 TCTGCTTGAACATCTACCACTCACTGAGTGAAGTGCAGCAGCTGCTGAGGCTGGA 60
OY 837 GTTGAGTGTGCTGGAAGGCGAGTCAAGATGACTTGTGTGATGATGACCGC 896
DB 61 GTTGAGTGTGCTGGAAGGCGAGTCAAGATGACTTGTGTGATGATGACCGC 120
OY 897 TGTGTGAACATGAGTGTGAGGCGAGTCAAGATGACTTGTGTGATGATGACCGC 956
DB 121 TGTGTGAACATGAGTGTGAGGCGAGTCAAGATGACTTGTGTGATGATGACCGC 180
OY 957 CACTGGGTTTAAGAAGATATCCAAACGTGTTTAAGAAGATCCGAGGCACTTGTGAAGAG 1016
DB 181 CACTGGGTTTAAGAAGATATCCAAACGTGTTTAAGAAGATCCGAGGCACTTGTGAAGAG 240
OY 1017 ACCGAGTGTGAGTGTGAGGCGAGTCAAGATGACTTGTGTGATGATGACCGC 1076
DB 241 ACCGAGTGTGAGTGTGAGGCGAGTCAAGATGACTTGTGTGATGATGACCGC 300
OY 1077 GCCATGAACATGATCTGTATCCAAACAGAGATGATGATGATGATGATGATGATGATGAT 1136
DB 301 GCCATGAACATGATCTGTATCCAAACAGAGATGATGATGATGATGATGATGATGATGAT 360
OY 1137 GAATCCATTTTGGATGAGGCTGGAAGAGACACAGATGTGATGTTGGTGGAAACAAGTG 1196
DB 361 GAATCCATTTTGGATGAGGCTGGAAGAGACACAGATGTGATGTTGGTGGAAACAAGTG 420
OY 1197 GTGGGTGTGAGTGTGAGTGTGAGGCGAGTCAAGATGACTTGTGTGATGATGACCGC 1256
DB 421 GTGGGTGTGAGTGTGAGTGTGAGGCGAGTCAAGATGACTTGTGTGATGATGACCGC 480
OY 1257 GCAATGTGTACATTAACCAAAATCCAGCCCATCTGTCTGTGAGGCGTCAATGAGTGG 1316
DB 481 GCAATGTGTACATTAACCAAAATCCAGCCCATCTGTCTGTGAGGCGTCAATGAGTGG 540
OY 1317 TTTGAGGTTGTAAGAGTGAAGTCAATCCGCAAGTGTGATGATGATGATGATGATGATGAT 1376
DB 541 TTTGAGGTTGTAAGAGTGAAGTCAATCCGCAAGTGTGATGATGATGATGATGATGATGAT 600
OY 1377 GGAATTAAGAATGTAGTGAACAGGAGCACTTGTGATGATGATGATGATGATGATGATGAT 1436
DB 601 GGAATTAAGAATGTAGTGAACAGGAGCACTTGTGATGATGATGATGATGATGATGATGAT 660
OY 1437 TGCATATATGGCCACT 1452
DB 661 TGCATATATGGCCACT 676

RESULT 12

BM547143

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM547143 1059 bp MRNA linear EST 20-FEB-2002
AGENCOURT_649536 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5730185
5', mRNA sequence.

BM547143
EST.

BM547143.1 GI:18780677

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1059)

NIH-MGC http://imgc.ncl.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgrabs-remail.nih.gov

Tissue Procurement: Invitrogen

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLM12728 row: k column: 18

High quality sequence stop: 747.

Location/Qualifiers

1. 1059

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5730185"

/clone_1db="NIH_MGC_124"

/tissue_type="hippocampus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMV-Sport6; Site: 1; EcoRV

(destroyed); Site: 2; NotI; RNA source male hippocampus,

age 27. Library is oligo-dT primed and directionally

cloned (EcoRV site is destroyed upon cloning). Average

insert size 1.4 kb, insert size range 0.9-4 kb. Library is

normalized and enriched for full-length clones and was

constructed by C. Gruber (Invitrogen). Research Genetics

tracking code 012."

BASE COUNT 290 a 242 c 221 g 304 t 2 others

ORIGIN

Query Match 26.4%; Score 676; DB 13; Length 1059;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1377 GGAATTAAGAATGTAGTGAACAGGAGCACTTGGATGCATGAAGAAACAGTTGATGTA 1436
DB 20 GGAATTAAGAATGTAGTGAACAGGAGCACTTGGATGCATGAAGAAACAGTTGATGTA 79
OY 1437 TGCATATATGGCCACTCCAAACAGAAATGATGTGACACAGCTCCGACCTCGGAGCTG 1496
DB 80 TGCATATATGGCCACTCCAAACAGAAATGATGTGACACAGCTCCGACCTCGGAGCTG 139
OY 1497 ACGTGGAGGAGTACCTTCTCAGGTGACCATCTCATCTGCGCAAGATGGCAACAGACT 1556
DB 140 ACGTGGAGGAGTACCTTCTCAGGTGACCATCTCATCTGCGCAAGATGGCAACAGACT 199
OY 1557 GTCTCTCTGAGAGGAGTGTCTACTCAATTTGAGTGTGCTCAGAGTTCCACCTTTGT 1616
DB 200 GTCTCTCTGAGAGGAGTGTCTACTCAATTTGAGTGTGCTCAGAGTTCCACCTTTGT 259
OY 1617 CTGTCCATCAACAGCACACACAGAGCTTGGACACTGTATGAAGTCTTAATATGACCCGAG 1676
DB 260 CTGTCCATCAACAGCACACACAGAGCTTGGACACTGTATGAAGTCTTAATATGACCCGAG 319
OY 1677 GGGGATACAGCAGAGTGTACTTGTCTTAAGAAAGATGAATAGTGTCCAGC 1736

Db 320 GGGGATACAGAGAGATGTGTTCTTCCATAGAAAATGATGATGCTTGGCAGC 379
 QY 1737 TTGCATCTCCATCATTTGATGGCCACCTTACAGAGCTGACAGATGACCAAGAAATAT 1796
 Db 380 TTGCATCTCCATCATTTGATGGCCACCTTACAGAGCTGACAGATGACCAAGAAATAT 439
 QY 1797 CTGGGAGCTAACAAAAATGGGCGATTCACCACTTAATTTATTCAGATCTATGAGACATA 1356
 Db 440 CTGGGAGCTAACAAAAATGGGCGATTCACCACTTAATTTATTCAGATCTATGAGACATA 499
 QY 1857 CTACCAAGAGACAGCTGACCTGACACACACACTTAAGAAATATTTTAAAGATACTT 1916
 Db 500 CTACCAAGAGACAGCTGACCTGACACACACACTTAAGAAATATTTTAAAGATACTT 559
 QY 1917 TTATTTCTTCTTACTGCTTCTCTGATTTTTCCTAATTTTCAATCTGTTT 1976
 Db 560 TTATTTCTTCTTACTGCTTCTCTGATTTTTCCTAATTTTCAATCTGTTT 619
 QY 1977 CATTCTATATCCAAATTTGACAGACACAGAGAACTGCTTACAGGCTTAAATGA 2036
 Db 620 CATTCTATATCCAAATTTGACAGACACAGAGAACTGCTTACAGGCTTAAATGA 679
 QY 2037 AATAGAGTTCAAGGT 2052
 Db 680 AATAGAGTTCAAGGT 695

RESULT 13
 AL520704 891 bp mRNA 11near EST 13-FEB-2001
 LOCUS AL520704 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB002L22 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL520704
 VERSION AL520704.1 GI:12784197
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 891)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr; Web: www.genoscope.cns.fr.
 Location/Qualifiers
 1. 891
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DB002L22"
 /clone_1b="LTI_NFL004_NBC2"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifestech.com URL: http://fulllength.invitrogen.com" 2 others

BASE COUNT 194 a 238 c 299 g 158 t
 ORIGIN
 Query Match 26.3%; Score 673; DB 9; Length 891;
 Best Local Similarity 99.5%; Pred No. 0;
 Matches 873; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGGGGGGGCAAGTGGGAGCTGGAGCTGCTGCTTCTGTTCTCTTGGCCACCGTGCT 60
 Db 15 GGGGGGGGCAAGTGGGAGCTGGAGCTGCTGCTTCTGTTCTCTTGGCCACCGTGCT 74
 QY 61 GTCCGGGTGCTTGGGCTGGCGAAGACAAAGGCTGGCCACAGCACTTGAAGCGGA 120
 Db 75 GTCCGGGTGCTTGGGCTGGCGAAGACAAAGGCTGGCCACAGCACTTGAAGCGGA 134
 QY 121 CGCAGCTCGACGAGGGGCGGCGAGAGGCTGCGATTCGCTGCGAGGGCGCCGCGC 180
 Db 135 CGCAGCTCGACGAGGGGCGGCGAGAGGCTGCGATTCGCTGCGAGGGCGCCGCGC 194
 QY 181 GGGGAGCGGGGCGGCGGCGAGAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 Db 195 GGGGAGCGGGGCGGCGGCGAGAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 254
 QY 241 GGGGCGGGGCGGAGATGCTGATGCTGACGCGATGCGCGCTCCGGGGTGGGGAGAGC 300
 Db 255 GGGGCGGGGCGGAGATGCTGATGCTGACGCGATGCGCGCTCCGGGGTGGGGAGAGC 314
 QY 301 TGAAGCAGGCGAAGAGATGAGAGAGCGCCGAGAGTACTCCTTCATGCGCCAGTCA 360
 Db 315 TGAAGCAGGCGAAGAGATGAGAGAGCGCCGAGAGTACTCCTTCATGCGCCAGTCA 374
 QY 361 AGCGGCCCAAGAGCAATTCAGTTGCTGATGACATGACAGAGTTTACCAATTCCTCA 420
 Db 375 AGCGGCCCAAGAGCAATTCAGTTGCTGATGACATGACAGAGTTTACCAATTCCTCA 434
 QY 421 CCAAACTGGCCGAGATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Db 435 CCAAACTGGCCGAGATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494
 QY 481 GTTCAGCTGATCTTACAGATAGCTGATGATGATGATGATGATGATGATGATGATGAT 540
 Db 495 GTTCAGCTGATCTTACAGATAGCTGATGATGATGATGATGATGATGATGATGATGAT 554
 QY 541 AAACCAACTCCAAAGGCGAGAGCAATTTCTGTGAAGAACATCAAGCAGGCGAGATTG 600
 Db 555 AAACCAACTCCAAAGGCGAGAGCAATTTCTGTGAAGAACATCAAGCAGGCGAGATTG 614
 QY 601 GAGCGCGGAGATTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 Db 615 GAGCGCGGAGATTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
 QY 661 GTCTCAGGGGAG 720
 Db 675 GTCTCAGGGGAG 734
 QY 721 CCCAGACAGCGGTGTTGATGAGACACTGTGCTGGGGGCTCAAGTCCGCTGCTG 780
 Db 735 CCCAGACAGCGGTGTTGATGAGACACTGTGCTGGGGGCTCAAGTCCGCTGCTG 794
 QY 781 CTGTGAACATCTACTCAACAGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 840
 Db 795 CTGTGAACATCTACTCAACAGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 854
 QY 841 CAGTGTGCTTGAAGGCGAGTCAAGATGACTT 877
 Db 855 CAGTGTGCTTGAAGGCGAGTCAAGATGACTT 891

RESULT 14
 BG708016 871 bp mRNA 11near EST 07-MAY-2001
 LOCUS BG708016 6026713A1F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793770.5,
 DEFINITION mRNA sequence.
 ACCESSION BG708016
 VERSION BG708016.1 GI:13984941
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

QY	661	GTGCTCAGGGGAGAACCCCTTGCGTGGGAGCTAAATATGCGGGCTGTACACATCAACAG	720
Db	1	GTGCTCAGGGGAGAACCCCTTGCGTGGGAGCTAAATATGCGGGCTGTACACATCAACAG	60
QY	721	CCGACAGAGCGGTGTGATTTAGACACACTGTGCCCCCTGGGGGCTCAGTGCCTCGTCTG	780
Db	61	CCGACAGAGCGGTGTGATTTAGACACACTGTGCCCCCTGGGGGCTCAGTGCCTCGTCTG	120
QY	781	CTTGTACATCTACTCAACTCAGAAATGAATGATGCTACACACTGGCTGAGGCTGAGATTG	840
Db	121	CTTGTACATCTACTCAACTCAGAAATGAATGATGCTACACACTGGCTGAGGCTGAGATTG	180
QY	841	CAGTGTTCGCTTGGAGAGCGGAGCTCAGAAAGATACTCTCGTGGTGTATTGACCGCTTG	900
Db	181	CAGTGTTCGCTTGGAGAGCGGAGCTCAGAAAGATACTCTCGTGGTGTATTGACCGCTTG	240
QY	901	TGAACATGATGGGTGGCGAGGCCAACATGATCCTGGATGATGGGGAGACTTAACCCACT	960
Db	241	TGAACATGATGGGTGGCGAGGCCAACATGATCCTGGATGATGGGGAGACTTAACCCACT	300
QY	961	GGGTTTAAAGAAAGTATCCAAAGGTGTTTAAAGAAATCCGAGGCAATGTGGAAAGAGACG	1020
Db	301	GGGTTTAAAGAAAGTATCCAAAGGTGTTTAAAGAAATCCGAGGCAATGTGGAAAGAGACG	360
QY	1021	TGACTGGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGAGCTGTGTCTCCGGGCA	1080
Db	361	TGACTGGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGAGCTGTGTCTCCGGGCA	420
QY	1081	TGAACGTCATGATTCGTGTTACCAAAACAGAGTTTATATTACTTGTATCTGTCGCGAGAT	1140
Db	421	TGAACGTCATGATTCGTGTTACCAAAACAGAGTTTATATTACTTGTATCTGTCGCGAGAT	480
QY	1141	CCATTGTGATGGGCTCAGAGGACCAACAGATGTGATTTGGTGGGAAACAAGGAGGG	1200
Db	481	CCATTGTGATGGGCTCAGAGGACCAACAGATGTGATTTGGTGGGAAACAAGGAGGG	540
QY	1201	TGCTGGCTATGGTAGGTAGGCAAGGGCTGTGTGCTCTCTCAAAGCTCTTTGGAGCA	1260
Db	541	TGCTGGCTATGGTAGGTAGGCAAGGGCTGTGTGCTCTCTCAAAGCTCTTTGGAGCA	600
QY	1261	TTGTGTCATATTACGAAATGAGCCCATGTGTGCTCTGAGGCGCTGCATGATGGG	1316
Db	601	TTGTGTCATATTACGAAATGAGCCCATGTGTGCTCTGAGGCGCTGCATGATGGG	656

Search completed: April 20, 2003, 22:06:43
Job time : 3278 secs

Db	121	CGCAGCTCGACGACGAGGGGCCGCGAGAGAGGGGTGGCGATCGCTGTCTGGAGGGGCCCGCC	180
Qy	181	GGCGAGGCGGGCGCGCGCCGACGAGGGGAAAGAGGGCGGGCGCGCGGTCAACCGCTGGCC	240
Db	181	GGCGAGGCGGGCGCGCGCCGACGAGGGGAAAGAGGGCGGGCGCGCGGTCAACCGCTGGCC	240
Qy	241	GGCGCGGGCGGGGAAATGTGATGTCTGTACGCGCATCGCTGCCCGGGGTGGGAGAGAC	300
Db	241	GGCGCGGGCGGGGAAATGTGATGTCTGTACGCGCATCGCTGCCCGGGGTGGGAGAGAC	300
Qy	301	TGAAGCAGGCGCAAGAGATGAGAGAGCGCCGAGAAAGTACTCTTCACTGGCCACCGTCACCA	360
Db	301	TGAAGCAGGCGCAAGAGATGAGAGAGCGCCGAGAAAGTACTCTTCACTGGCCACCGTCACCA	360
Qy	361	AGGCGCCCAAGAAAGCAAAATCCAGTTTGTGTATGACATGCAAGGATTCACCAAAATTCCCA	420
Db	361	AGGCGCCCAAGAAAGCAAAATCCAGTTTGTGTATGACATGCAAGGATTCACCAAAATTCCCA	420
Qy	421	CCAAATATGGCCGAAAGATCTTTGTCTGTCCGTGATCTCACAGTCCCTCCACTGACAGCTACA	480
Db	421	CCAAATATGGCCGAAAGATCTTTGTCTGTCCGTGATCTCACAGTCCCTCCACTGACAGCTACA	480
Qy	481	GTTCAGCTGATCTCTACACAGATAGCTGTGATGATGAGTTTCTCCCGAGAGAAAGAC	540
Db	481	GTTCAGCTGATCTCTACACAGATAGCTGTGATGATGAGTTTCTCCCGAGAGAAAGAC	540
Qy	541	AAACCAACTCCAAAGGCGACGACGCAATTTCTGTGTGAAGAAATCAAGACGCGAGAAATTG	600
Db	541	AAACCAACTCCAAAGGCGACGACGCAATTTCTGTGTGAAGAAATCAAGACGCGAGAAATTG	600
Qy	601	GACCGCCGGGAGATTGAGATTGCGAGACAGACATGTCTGTCTGTATTCATCTACAGAAAC	660
Db	601	GACCGCCGGGAGATTGAGATTGCGAGACAGACATGTCTGTCTGTATTCATCTACAGAAAC	660
Qy	661	GTGCTCAGGGGAGAAAGCCCTTGGCTGTGTCTAAATATAGTGGGCTGTACACATCAGAC	720
Db	661	GTGCTCAGGGGAGAAAGCCCTTGGCTGTGTCTAAATATAGTGGGCTGTACACATCAGAC	720
Qy	721	CCGACAGCGGGGTGTGATTGAGACACTGTGTGCCCTGGGGCTCAGTCCGCTGTGTCTG	780
Db	721	CCGACAGCGGGGTGTGATTGAGACACTGTGTGCCCTGGGGCTCAGTCCGCTGTGTCTG	780
Qy	781	CTTGTACATCTTCTCTCAACTCAGAAAGAGTATGCTGACAGCACTGGCTGATGGCTGGAGTGTG	840
Db	781	CTTGTACATCTTCTCTCAACTCAGAAAGAGTATGCTGACAGCACTGGCTGATGGCTGGAGTGTG	840
Qy	841	CAGTGTCCGCTTGGAGGGGAGTCTCAGAGATGACTCTGGGTGTATGTACCGCTGTG	900
Db	841	CAGTGTCCGCTTGGAGGGGAGTCTCAGAGATGACTCTGGGTGTATGTATGCCGCTGTG	900
Qy	901	TGAACATGATGGGTGGCAGGCCCAACATGATCCTGGATGATGGGGAGAACTTAAACCACT	960
Db	901	TGAACATGATGGGTGGCAGGCCCAACATGATCCTGGATGATGGGGAGAACTTAAACCACT	960
Qy	961	GGGTTTAAAGAAATATCAAAAGTGTTTAAGAAAGATCCGAGGCAATGTGGAAGAGACG	1020
Db	961	GGGTTTAAAGAAATATCAAAAGTGTTTAAGAAAGATCCGAGGCAATGTGGAAGAGACG	1020
Qy	1021	TGACTGGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGGCA	1080
Db	1021	TGAGTGGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGCTCTGTGTCCGGCA	1080
Qy	1081	TGAAGCAGTGAATGTGTGTACCAAAACAGAAAGTTGATTAATCTGTACGTGTCCGAGAT	1140
Db	1081	TGAAGCAGTGAATGTGTGTACCAAAACAGAAAGTTGATTAATCTGTACGTGTCCGAGAT	1140
Qy	1141	CCATTTGGATGGCTTGAAGAGGACCAACAGATGTATGTTTGTGGGAAACAAAGTGTGG	1200
Db	1141	CCATTTGGATGGCTTGAAGAGGACCAACAGATGTATGTTTGTGGGAAACAAAGTGTGG	1200
Qy	1201	TGTGTGGCTATGGTGAAGTATGCGACAGGCTGTGTGTCTCTCAAGCTCTTGGAGCA	1260
Db	1201	TGTGTGGCTATGGTGAAGTATGCGACAGGCTGTGTGTCTCTCAAGCTCTTGGAGCA	1260

Dd	1201	TGCTGGCTATG	TGAGTG	TAGGCA	AGGGCGT	CGTGTG	GTGCTG	CTCA	AAAGCT	TTTG	AGACA	1280
Qy	1261	TTGTCTAC	ATTACG	GAATG	AGACCC	CACTGT	GTCTG	TCG	AGGCG	CTG	AGATGG	1320
Dd	1261	TTGTCTAC	ATTACG	GAATG	AGACCC	CACTGT	GTCTG	TCG	AGGCG	CTG	AGATGG	1320
Qy	1321	GGGTGGTAA	AGCTAA	TGAATG	ATCCG	GAAC	TGCTG	ATAT	CACTT	GC	ACAGAA	1380
Dd	1321	GGGTGGTAA	AGCTAA	TGAATG	ATCCG	GAAC	TGCTG	ATAT	CACTT	GC	ACAGAA	1380
Qy	1381	ATPAGATAT	GTAG	AGAC	AGGAG	CAC	CTTG	ATG	GCAT	GA	AAAAAC	1440
Dd	1381	ATPAGATAT	GTAG	AGAC	AGGAG	CAC	CTTG	ATG	GCAT	GA	AAAAAC	1440
Qy	1441	ATATGGGCA	CTCCAA	CAC	GA	AAATG	ATG	ATG	AC	AGCT	CCGAG	1500
Dd	1441	ATATGGGCA	CTCCAA	CAC	GA	AAATG	ATG	ATG	AC	AGCT	CCGAG	1500
Qy	1501	GGGAGGAG	ATAG	CTTCT	CA	GGTGG	AG	CA	ATG	CTG	GGCA	1560
Dd	1501	GGGAGGAG	ATAG	CTTCT	CA	GGTGG	AG	CA	ATG	CTG	GGCA	1560
Qy	1561	TCCTGGCAG	AGGGT	CCGT	CT	ACTCA	ATTG	AG	CGTCC	AC	AGTTCC	1620
Dd	1561	TCCTGGCAG	AGGGT	CCGT	CT	ACTCA	ATTG	AG	CGTCC	AC	AGTTCC	1620
Qy	1621	CCATCACAG	CCAC	AC	AC	AG	CGTT	TGG	CA	TAG	AACT	1680
Dd	1621	CCATCACAG	CCAC	AC	AC	AG	CGTT	TGG	CA	TAG	AACT	1680
Qy	1681	GATACAG	AGAG	ATG	ATG	CTT	CT	CT	CT	CT	CT	1740
Dd	1681	GATACAG	AGAG	ATG	ATG	CTT	CT	CT	CT	CT	CT	1740
Qy	1741	ATCTGCAT	CA	TAT	TG	ATG	CC	CA	CTT	AC	AG	1800
Dd	1741	ATCTGCAT	CA	TAT	TG	ATG	CC	CA	CTT	AC	AG	1800
Qy	1801	GACTCA	CA	CA	AA	AA	TGG	CC	AT	T	CA	1860
Dd	1801	GACTCA	CA	CA	AA	AA	TGG	CC	AT	T	CA	1860
Qy	1861	CAAGSAC	AG	CT	CA	AG	CA	CA	CA	CT	CA	1920
Dd	1861	CAAGSAC	AG	CT	CA	AG	CA	CA	CA	CT	CA	1920
Qy	1921	TTTTCTT	CT	ACT	CT	T	CC	CT	T	GA	TT	1980
Dd	1921	TTTTCTT	CT	ACT	CT	T	CC	CT	T	GA	TT	1980
Qy	1981	TCATAT	AT	CA	AG	TT	TC	GA	CA	CA	CA	2040
Dd	1981	TCATAT	AT	CA	AG	TT	TC	GA	CA	CA	CA	2040
Qy	2041	GAA	GT	CA	AG	GG	T	CC	CA	CT	CA	2100
Dd	2041	GAA	GT	CA	AG	GG	T	CC	CA	CT	CA	2100
Qy	2101	GGG	AT	TC	CT	TC	TT	CA	CTT	CT	GA	2160
Dd	2101	GGG	AT	TC	CT	TC	TT	CA	CTT	CT	GA	2160
Qy	2161	GGA	AT	TC	CT	TC	TT	CA	CTT	CT	GA	2220
Dd	2161	GGA	AT	TC	CT	TC	TT	CA	CTT	CT	GA	2220
Qy	2221	GGC	AT	TC	CT	TC	TT	CA	CTT	CT	GA	2280
Dd	2221	GGC	AT	TC	CT	TC	TT	CA	CTT	CT	GA	2280
Qy	2281	ACC	T	AC	CC	GG	CA	GG	AT	GG	AT	2340
Dd	2281	ACC	T	AC	CC	GG	CA	GG	AT	GG	AT	2340
Qy	2340	ACCT	AC	CC	GG	CA	GG	AT	GG	AT	GG	2400
Dd	2340	ACCT	AC	CC	GG	CA	GG	AT	GG	AT	GG	2400

OY	1322	GGTGGTAAAGCTAAATGAAGTCATCCGGCCAGAGTCGTAATTAAGTTCGACAGAGAA	1381
Db	1	GGTGGTAAAGCTAAATGAAGTCATCCGGCCAGAGTCGTAATTAAGTTCGACAGAGAA	60
OY	1382	TAAAGATGTAGTCAGCAGCGGAGCACTTGGATCGCATGAAAAAACAGTTGTATGTGCAA	1441
Db	61	TAAAGATGTAGTCAGCAGCGGAGCACTTGGATCGCATGAAAAAACAGTTGTATGTGCAA	120
OY	1442	TATGGGCGCACTCCAAACACAGAAATCGATGTAGACACCTCCGCACTCCGAGACTACAGTG	1501
Db	121	TATGGGCGCACTCCAAACACAGAAATCGATGTAGACACCTCCGCACTCCGAGACTACAGTG	180
OY	1502	GGACGAGTAGTCGTCAGAGTGGACCATCTCATCTGGCCAGATGGGAAACGAGTTGTCT	1561
Db	181	GGACGAGTAGTCCTCAGGTGGACCATCTCATCTGGCCAGAT - GAAACGAGTTGTCT	239
OY	1562	CTGTGACAGAGGCGTCTACTCAATTTGAGCTGCTCCACAGTCCCATCTTGTCTCTC	1621
Db	240	CTGTGACAGAGGCGTCTACTCAATTTGAGCTGCTCCACAGTCCCATCTTGTCTCTC	299

Oy	1622	CATACAGCCCAACACAGGCTTTGGCACTGATAGAACTCTTAATGACACCCGAGGGCG	1681
Db	300	CATACAGCCCAACACAGGCTTTGGCACTGATAGAACTCTTAATGACACCCGAGGGCG	359
Oy	1662	ATACAGCAGGATGTGTACTTGTCTCTAAGAAA--TGAGTGAATACGTTGGCAGCTTGC	1740
Db	360	ATACAGCAGGATGTGTACTTGTCTCTAAGAAAATTTGATGAATACGTTGGCAGCTTGC	419
Oy	1741	ATTCGCATCATTTGGATGGCCACCTTACAGAGCTGACATGACCAAGAAAATTAATCTGG	1800
Db	420	ATTCGCATCATTTGGATGGCCACCTTACAGAGCTGACATGACCAAGC-AAATATCTGG	478
Oy	1801	GACTCAACAAAAATGGGCACTTCAAACTAATTATTATACAGATCTAATGGAACATATAC	1860
Db	479	GACTCAACAAAA--TGGGCCATTCATCAACTAATTATTATACAGATCTAATGGAACATATAC	536
Oy	1861	CAAGGACCAGTCCAC	1875
Db	537	CAAGGACCAGTCCAC	551

```

RESULT 4
US-09-879-536-850/c
: Sequence 850. Application US/0979536
: Patent No. US20020144298A1
: GENERAL INFORMATION:
: APPLICANT: Endeavor, Wilson O.
: APPLICANT: Steinmann, Kathleen E.
: APPLICANT: Astle, Jon H.
: APPLICANT: Burgess, Christopher C.
: APPLICANT: Bushnell, Steven E.
: APPLICANT: Carroll, Irl Eddie
: APPLICANT: Catino, Theodore J.
: APPLICANT: Dertl, Adnan
: APPLICANT: Ford, Donna M.
: APPLICANT: Lewis, Marcia E.
: APPLICANT: Monahan, John E.
: APPLICANT: Schlegel, Robert
: TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
: TITLE OF INVENTION: PRODUCTS
: FILE REFERENCE: CCD-257 (US)
: CURRENT APPLICATION NUMBER: US/09/879,536
: CURRENT FILING DATE: 2001-09-21
: PRIOR APPLICATION NUMBER: US 60/088,801
: PRIOR FILING DATE: 1998-06-10
: NUMBER OF SEQ ID NOS: 850
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 850
: LENGTH: 636
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(636)
: OTHER INFORMATION: n = A,T,C or G
US-09-879-536-850

```

Query Match	15.9%	Score 406.4	DB 10	Length 636
Best Local Similarity	99.5%	Pred. No. 3.6e-105		
Matches 418	Conservative 0	Mismatches 1	Indels 1	Gaps 1
QY	707	TACACACATCATCAGAGCCAGACAGC-GGTGTTGATGAGACACTGTGCGCCGTGGGGGCTC	765	
Db	420	TACACACATCATCAGAGCCAGACAGCGGGGTGTGATTGAGACACTGTGTCTGTGGGGGCTC	361	
QY	766	AGTGCAGCTGTCTCTCTTTAACTACTCTCACTCAGAAATGAAGTAGCTGCAGCACTGG	825	
Db	360	AGTGCAGCTGTCTCTTTAACTACTCTCACTCAGAAATGAAGTAGCTGCAGCACTGG	301	
QY	826	CTGAGGCTGGAATTCACATGTTCCTGTTGGAAAGGCGAGTCAGAAAGTAGTCTTGTGTGT	885	
Db	300	CTGAGGCTGGAATTCACATGTTCCTGTTGGAAAGGCGAGTCAGAAAGTAGTCTTGTGTGT	241	

QY	886	GTGTGACCGCTGTGTAAATATGGATGGGTCGACAGGCAAAATGATCCGTGAATGATG	945
Db	240	GATATGACCGCTGTGTAAATATGGATGGGTCGACAGGCAAAATGATCCGTGAATGATG	181
QY	946	GAGACTTAACCCACTGGGTATTATAGAGATNTCCAAAGCTGTTTAAAGATCCAGGCA	1005
Db	180	GAGACTTAACCCACTGGGTATTATAGAGATNTCCAAAGCTGTTTAAAGATCCAGGCA	121
QY	1006	TTGTGGAAGAGAGCGTACGTGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAC	1065
Db	120	TTGTGGAAGAGAGCGTACGTGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAC	61
QY	1066	TCTGTGTCGGGCATGACATGCTAAAGATATCTTTACCAAGACGAGAGTTTGATTAATCTGT	1126
Db	60	TCTGTGTCGGGCATGACATGCTAAAGATATCTTTACCAAGACGAGAGTTTGATTAATCTGT	1

```

RESULT 5
US-09-925-301-217
: Sequence 217, Application US/09925301
: Patent No. US20020052308A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA106
: CURRENT APPLICATION NUMBER: US/09/925,301
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05882
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 1694
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 217
: LENGTH: 2200
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (2188)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-217

```

Query Match	15.8%	Score 404	DB 10	Length 2200
Best Local Similarity	58.4%	Pred. No. 3.9e+04		
Matches 745	Conservative	0	Mismatches 525	Indels 6
				Gaps 2
QY	580	ACATCAAGCAGGACCAATTTGGACCCCGGAGATTTGAGATTCTCAGAGCCAAACATGTCTG	639	
Db	108	ACATGGGCGCTGGCTGCTCCCTGGGGAGCCAAAGCCCTTGACATTTCTTGAGAACGAGATGCCGG	167	
QY	640	CTCTGATTTCACCTACGAGAAACGTGCTCAAGGGGGAGAAAGCCCTTGGCTGGTCTAAATATG	699	
Db	158	GCGTATATCGGATATGGGGGAGCGGTACTGGCCCTCCAAAGCCACTGAAGGCGCCCGCATCG	227	
QY	700	TGGCGCTGACACACATCATCACAGCCAGACAGCGGTGTTGATTGAGACACTCTGTGCCCTGG	759	
Db	228	CTGGCTGCCCTGCACATGACCGGAGAGAGGGCCGCTCTCATTTAGACCTCTGCACACCTGG	287	
QY	760	GGGCTCAGTGGCCGCTGCTGCTGTTTAACTACTACTCAACTCAGATGAGATAGCTGCG	819	
Db	288	GTCGTGAGGTGCAAGTGGTCCAGCTCACAACATTTTCTCCACCAGGACCAATGCGGGGGCTG	347	
QY	820	CACCTGCTGAGGCTGGAGATTGCATGTTCCGCTTGGAGGGCGAGTCAGAAAGATGACTTCT	879	
Db	348	CCATTCGCCAAGGCTGGCGCATTTCCGGTGTATGCTCTGGAAGGGCGCAAAAGGACGAGAGTACC	407	
QY	880	GGTGTGATTATGACCGCTGTGTGAACATGATGGGTGGCAGGGCCAAACATGATCCTGGATG	939	
Db	408	TGTGTGCAATTGAGAGAGACCCTGTACTCTTCAGAGAGGGGCCCTCAACATGATTTGAGAG	467	
QY	940	ATGGGGGAGACTTAAACCACCTGGGTTTATAGAAAGTATCCAAACGCTGTTTAAAGAATCC	999	

Db	468	ACGGGGGGGACCTCCACCAACCTCATCCACCAAGTACCGGACGCTCTGCGCAGGACATCC	1111	527
QY	1000	GAGGATCTGTGGAAAGAGACCGTGTACTGTGTGTACAGGCTGTATCACTCTCCAAAGTC	1111	1059
Db	528	GAGGATCTGTGTAGGAGACACAGACTGGGGTCCACAACTCTACMAATGATGGCCAAATG	1111	587
QY	1060	GGAAAGCTCTGTGTCCGGCCCATGAACGTCAATGATCTGTATACCAAGAGATTGATA	1111	1119
Db	588	GGATCTCTCAAGGTGCTCGCATCTCAATGTCAATGATCTCGTACCAAGAGAGATTGATA	1111	647
QY	1120	ACTGTACTGCTGCCGAGATCCATTTTGGATGGCCCTGAAGAGGACACAGATGTATGT	1111	1179
Db	648	ACCTTATGTGGCTCGCGGAGATCCCTCATAGATGGATCAAGCGGCGACAAATGTGATGA	1111	707
QY	1180	TTGGTGGGAAACAAGTGGGTGTGTGTGGCATGTGTGAGTGAAGCAAGGGCTGTGTGTGT	1111	1239
Db	708	TTGGCGGCAAGGTAGGGGTGTGTGTGACAGGCTGTGTGTGTGTGTGGGCAAGGGCTGTGTGTGT	1111	767
QY	1240	CTCTCAAGAGCTTTGGAGCAATTTGTCTACATTAACCGAAATCGAACCCCATCTGTGCTGTGC	1111	1299
Db	768	CCCTCGGGGGTTTGGAGGCCCGCGTCAATCAACCGAGATGTAGGCCCATCAACGCATCTGC	1111	827
QY	1300	AGGCTGCATGATGGGTTCAGGGGTGTAAAGCTAAATGAATCATCCGCGAATGTGATG	1111	1359
Db	828	AGGCTGCCATGTAGAGGGCTGTAGAGTGAACCCATGTAGTGAAGGCTGTGAGAGGGCAACA	1111	887
QY	1360	TCTGTAATACTGTGCACAGGAATAATGAATGTATGACACGGGACGACTGTGATGGCATGA	1111	1419
Db	888	TCTTTGTGTCACACACAGGCTGTATTGACATCATCTTTGGCGGCGACTTTGAGCAGATGA	1111	947
QY	1420	AAACAGTTGTATCTGTATGCATATATGAGGCGCATCTCCAAACAGAAATGATGTGACACGCC	1111	1479
Db	948	AGGATGATGCCATTTGTGTGTATTAACATTTGAGACACTTTGACGTGGAGATGATGTCAAGTGGC	1111	1007
QY	1480	TCCGCACCTCCGAGAGCTGACGTGGGAGGAGTACGTTTTCAGGTGGACATGTCTATCTGGC	1111	1539
Db	1008	TCAACGAGAACCCCGTGGAGAAAGGTGAMCATCAAGCCGAGGTGAGCCGATTCGGTTGA	1111	1067
QY	1540	CAGATGGCAAAAGAGTTGTACCTCTGTGGACAGAGGTCGTCTACTCAATTTGAGTGCCTCA	1111	1599
Db	1068	AGAAATGGGCGCGCATCATCTCTGTGGCCGAGGTCGAGCTGTGTAACTGGGTGTGTGCA	1111	1127
QY	1600	CAG---TTCCACACTTTGTGTGTGTCCATCAACAGCCACACACAGAGCTTTGGCACTGATAG	1111	1656
Db	1128	TGGGCGCACCCAGCTCGATGAGTAACTCTTACACCAACAGAGTGTATGCGCAGATCG	1111	1187
QY	1657	AACTCTATATATGACACCCCGAGGGCGCATACAGCAGATGTGTACTGTCTTCTTAAGAAA	1111	1716
Db	1188	AGCTGTGGACCCATCCAGACAAAG--TACCCGTTGGGGTTCATTTCTGTGCCAAGAAAGC	1111	1244
QY	1717	TGGAGTAATACGTTGGCACCTTGCATCTGCTCATCATTTGATAGCCACCTTACAGAGCTGA	1111	1776
Db	1245	TGGATGAGGCAATGGCTGAAAGCCCACTGTGGCAAGCTGATGTGAAGTGAACAAAGCTAA	1111	1304
QY	1777	CAGATGACCAAGCAAAATATCTGTGGACTCAACAAAAATGGGCAATTCAAACCTAATTAAT	1111	1836
Db	1305	CTGAAGAAGCAACCCAGTACTGTGGCATGTCTGTGTATGGGCCCTTCAAGCCGATCACT	1111	1364
QY	1837	ACAGATACATTAATGAC 1852		
Db	1365	ACGCTACTGAGAGCC 1380		

RESULT 6
 US-10-044-090-344
 ; Sequence 344, Application US/10044090
 ; Patent No. US20020137081A1
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Olga Bandman
 ; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
 ; FILE REFERENCE: PA-0028 US
 ; CURRENT APPLICATION NUMBER: US/10/044, 090

```

1 CURRENT FILING DATE: 2002-01-09
2
3 NUMBER OF SEQ ID NOS: 850
4
5 SOFTWARE: PERL Program
6
7 SEQ ID NO 344
8
9 LENGTH: 2429
10
11 TYPE: DNA
12
13 ORGANISM: Homo sapiens
14
15 FEATURE:
16
17 NAME/KEY: misc_feature
18
19 OTHER INFORMATION: incyte ID NO. US20020137081a1 1468237CB1
20
21 US-10-044-030-344

```

Query Match	Similarity	58.4%	Pred. No. 4.2e-104;	Length 2429;
Matches 745;	Conservative	0;	Mismatches 525;	Indels 6;
				Gaps 2;
QY	580	ACATATACGAGGACGAAATTTGGACGCCCGGAGATTGAGATTGGACAGCAAGCATGTGTG	639	
DB	108	ACATATACGAGGACGAAATTTGGACGCCCGGAGATTGAGATTGGACAGCAAGCATGTGTG	167	
QY	640	CTCTGATTTTCACTACAGGAAAGCTGCTCAAGGGGGAGAGCCCTTGGCTGTCTAAATATG	699	
DB	168	GCCCTATCGCTGATGCGGGAGCGGTACTCGCCCTCCAAAGCATGTAAGGCGCCGCCATCG	227	
QY	700	TGGGCTGTACACATACACAGCCGACAGCGCGTGTGATGTAGACACTGTGTGCGCTGG	759	
DB	228	CTGGCTGTACACATACACAGCCGACAGCGCGTGTGATGTAGACACTGTGTGCGCTGG	287	
QY	760	GAGGCTCAATGCGCGTGTGCTGTCTTAACATCTACTCACTCAGATGAAATGAGCTGAG	819	
DB	288	GTCGTGAGGTGAGTGTGCTGTCTTAACATCTACTCACTCAGATGAAATGAGCTGAG	347	
QY	820	CACGTGGAGGCTGAGGAGTGTGCTGTGCTTGGAGGGCGAGTCAGAAATGACTTCT	879	
DB	348	CCATGCGCAAGGCTGAGGAGTGTGCTGTGCTTGGAGGGCGAGTCAGAAATGACTTCT	407	
QY	880	GGTGTGTATTTGACCGCTGTGTGAACATGATGATGATGATGATGATGATGATGATG	939	
DB	408	TGTGTGTATTTGACCGCTGTGTGAACATGATGATGATGATGATGATGATGATGATG	467	
QY	940	ATGGGAGGAGACTTAAACCACTGGGTATTAAAGAAATATCAACCTGTATTAAAGAAATC	999	
DB	468	ACGGGGGAGGAGCTTAAACCACTGGGTATTAAAGAAATATCAACCTGTATTAAAGAAATC	527	
QY	1000	GAGGCAATTTGGAGAGAGAGCTGATGTGTTCACAGGCTGTATCATGCTTCCAAAGTGT	1059	
DB	528	GAGGCAATTTGGAGAGAGAGCTGATGTGTTCACAGGCTGTATCATGCTTCCAAAGTGT	587	
QY	1060	GGAACCTCTGTGTTCGCCGACATGAACATGATGATGATGATGATGATGATGATGATG	1119	
DB	588	GGAACCTCTGTGTTCGCCGACATGAACATGATGATGATGATGATGATGATGATGATG	647	
QY	1120	ACTTGTACTGTGCGCGAGAAATCAATTTTGGATGAGGCTCAAGAGGACCAAGATGTGATG	1179	
DB	648	ACCTTATGTGCTGCGCGAGAAATCAATTTTGGATGAGGCTCAAGAGGACCAAGATGTGATG	707	
QY	1180	TTTGTGGGAAACAGTG	1239	
DB	708	TTTGTGGGAAACAGTG	767	
QY	1240	CTCTCAAAAGCTTGTGGACATATGCTCTCATATTAACGAAATGACCCCATCTGTCTGCG	1299	
DB	768	CCCTCCGCGGCTTGTGGACATATGCTCTCATATTAACGAAATGACCCCATCTGTCTGCG	827	
QY	1300	AGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1359	
DB	828	AGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	887	
QY	1360	TGTGTATTAATCTTGAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG	1419	
DB	888	TGTGTATTAATCTTGAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG	947	
QY	1420	AAACAGTTGTATGCTATGACATATATGAGGACCTCCACACAGAAATGATGATGACACCC	1479	

	Query March	Best Local Similarity	94.2%;	Score 365;	DB 10;	Length 376;
	Matches	376	Conservative	0;	Mismatches	0; Indels 1; Gaps
OY	2174	TACCTTCTCTGTGGAAACATCTGCATCTGTCTTAATCGCTTTAAAGACCCATTTCCTAG	2233			
Db	376	TACCTTCTCTGTGGAAACATCTGCATCTGTCTTAATCGCTTTAAAGACCCATTTCCTAG	317			
OY	2234	CTGCTGAATCAGTGCCTCTTCACTCTTCACAGGAAGCAGGATGTACTACCGGCGAG	2293			
Db	316	CTGCTGAATCAGTGCCTCTTCACTCTTCACAGGAAGCAGGATGTACTACCGGCGAG	257			
OY	2294	GTACGTTAGATGTGGGTGTGCATGTTAATTCCTTAGAAGTCCAAAGCCCTGTTTCT	2353			

Query Match	Best Local Similarity	13.14%	Score 3.37	DB 10	Length 384
Matches 381	Conservative	0	Mismatches	0	Indels 4
					Gaps
QY 992	GAAATCCGAGGCAATTGTGGAAAGACGTCGACTGTGTTCACAGGCTGTATCAGCTCTC	105			
Db 1	GAAATCCGAGGCAATTGTGGAAAGACGTCGACTGTGTTCACAGGCTGTATCAGCTCTC	60			
QY 1052	CAAAAGCTGGGAAGCTCTGTGTTCGGGCGCATGAAAGCTAAATGATTCTCTTACCAAAACGAA	111			
Db 61	CAAAAGCTGGGAAG-CTCTGTGTTCGGGCGCATGAAAGCTAAATGATTCTCTTACCAAAACGAA	119			
QY 1112	GTTTGATPAACTGTACTAGTGCCTCCGAGAAATCCATTGTGGATGGCCCTGAAGAGGACACAGA	117			
Db 120	GTTTGATPAACTGTACTAGTGCCTCCGAGAAATCCATTGTGGATGGCCCTGAAGAGGACACAGA	179			
QY 1172	TGTGATGTTTGGTGGGGAACAAGTGTGTGTGTGTGGCTATGTGTAGGTAGGCAAGGGCTG	123			
Db 180	TGTGATGTTTGGTGGGGAACAAGTGTGTGTGTGTGGCTATGTGTAGGTAGGCAAGGGCTG	229			
QY 1232	-CTGTGCTGCTCTC-AAAGCTCTTGAAGCAATTGTCTACATTACCGAAATGACCCCATC	128			
Db 240	CCTGTGCTGCTCTCAAAAAGCTCTTGAAGCAATTGTCTACATTACCGAAATGACCCCATC	299			
QY 1290	TGTGCTCTGCAGAGGCTCGATGATGGGTTCAAGGTCGT-AAAGCTAAATGAATCATCCG	134			
Db 300	TGTGCTCTGCAGAGGCTCGATGATGGGTTCAAGGTCGTAAAGCTAAATGAATCATCCG	359			

OY 1349 GCAAGTCGATGCTGTAATACCTGC 1373
 |||||
 Db 360 GCAAGTCGATGCTGTAATACCTGC 364

RESULT 9
 US-09-884-441-40/c

; Sequence 40, Application US/09884441
 ; Patent No. US20020119158A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Algate, Paul A.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.462C7
 ; CURRENT APPLICATION NUMBER: US/09/884,441
 ; CURRENT FILING DATE: 2001-06-18
 ; NUMBER OF SEQ ID NOS: 489
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 40
 ; LENGTH: 292
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-884-441-40

Query Match 10.2%; Score 261.4; DB 10; Length 292;
 Best Local Similarity 99.3%; Pred. No. 3.9e-64;
 Matches 273; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 2261 TTCAGAGAGACGAGATGTGTACTACCCGACGAGTAGTAGTGTGTGTGCATGTT 2320
 |||||
 Db 282 TTCAGAGAGACGAGATGTGTACTACCCGACGAGTAGTAGTGTGTGTGCATGTT 223
 OY 2321 AATTCCCTTAAGATGCCAAGCCCTGTTCTGCGTAAAGGTGATGTCAGTTGAGA 2380
 |||||
 Db 222 AATTCCCTTAAGATGCCAAGCCCTGTTCTGCGTAAAGGTGATGTCAGTTGAGA 163
 OY 2381 GATGTGTAATAGAGATGCTTGTAAAGTACAGAGAGCCACTTGGATTATAGTAG 2440
 |||||
 Db 163 GATGTGTAATAGAGATGCTTGTAAAGTACAGAGAGCCACTTGGATTATAGTAG 103
 OY 2441 CCCTTCCTCCACTCCACGAGATGCTCATTTTTCAGATTATTAAGTACACTACTCT 2500
 |||||
 Db 102 CCCTTCCTCCACTCCACGAGATGCTCATTTTTCAGATTATTAAGTACACTACTCT 43
 OY 2501 AATTGAGTTAATTTTGTCTCTAGATTATTTTC 2535
 |||||
 Db 42 A-TTGAGTTAATTTTGTCTCTAGATTATTTTC 9

RESULT 10
 US-09-759-990-1

; Sequence 1, Application US/09759990
 ; Patent No. US20020119491A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Anticancer, Inc.
 ; APPLICANT: Xu, Mingxu
 ; APPLICANT: Han, Qiangdong
 ; TITLE OF INVENTION: HIGH EXPRESSION AND PRODUCTION OF HIGN
 ; TITLE OF INVENTION: SPECIFIC ACTIVITY RECOMBINANT S-ADENOSYLHOMOCYSTEINASE
 ; FILE REFERENCE: 31276-20026.00
 ; CURRENT APPLICATION NUMBER: US/09/759,990
 ; CURRENT FILING DATE: 2001-01-12
 ; PRIOR APPLICATION NUMBER: US 60/176,444
 ; PRIOR FILING DATE: 2000-01-14
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1461
 ; TYPE: DNA
 ; ORGANISM: Unknown

; FEATURE:
 ; OTHER INFORMATION: Nucleotide sequence encoding SAH
 US-09-759-990-1

Query Match 9.5%; Score 244; DB 10; Length 1461;
 Best Local Similarity 57.9%; Pred. No. 1e-58;
 Matches 493; Conservative 0; Mismatches 350; Indels 9; Gaps 3;

OY 1002 GGCATTGTGGAAGACGCGATGCTGCTACAGGGGTGATGCTGCTCAAGGCTGGG 1061
 |||||
 Db 613 GGTGTTCCGAGAGACAAACAGGTGCTCACCGCTCTACAGCTGAGAGAGAGGC 672
 OY 1062 AAGCTCTGTGTTCCGAGATGACGATGATGATGCTGTACCAAGAGATTGATGATAC 1121
 |||||
 Db 673 AAGCTCTGTGTTCCGAGATGACGATGATGATGCTGTACCAAGAGATTGATGATAC 732
 OY 1122 TTGACTGCTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1181
 |||||
 Db 733 ATCTACGCTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 792
 OY 1182 GGTGGAAGAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241
 |||||
 Db 793 GCGGCAAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 852
 OY 1242 CTCAGAGCTGTTGAGCAATGCTTCTACATACCGAAATCGACCCATGCTGCTGCTG 1301
 |||||
 Db 853 CTCGCTGCGCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912
 OY 1302 GCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361
 |||||
 Db 913 GCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 972
 OY 1362 GTATTAATCTGCAAGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1421
 |||||
 Db 973 TTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
 OY 1422 AACAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1481
 |||||
 Db 1033 GATAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1092
 OY 1482 CGCA---CTCCGAGCTGAGCTGAGGAGAGAGATGATGATGATGATGATGATGAT 1538
 |||||
 Db 1093 ATGAAATACCGAGCATCAAGCATCCATCAAGCCAGAAATGATGAGAAATTC 1152
 OY 1539 CCAGATGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1598
 |||||
 Db 1153 CCAGATGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
 OY 1599 ACAAGT---CCCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1655
 |||||
 Db 1213 ACAGGTACCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
 OY 1656 GAACTGTATATGACCCGAGGAGGAGATGATGATGATGATGATGATGATGATGAT 1715
 |||||
 Db 1273 GACCTCTACGA---AAGAGAGAGAAATCTCGAAGAGAGGTTTACACCTCCAGAGAT 1329
 OY 1716 ATGATTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1775
 |||||
 Db 1330 CTCGATTAAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1389
 OY 1776 ACAGATACCAAGAGAAATATCTGGAGCTCAACAAATGGCCATCAACATTAATAT 1835
 |||||
 Db 1390 ACACAGAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1449
 OY 1836 TACAGATATGAA 1847
 |||||
 Db 1450 TACCGTTATTA 1461

RESULT 11
 US-09-746-660A-97
 ; Sequence 97, Application US/09746660A
 ; Publication No. US20030049804A1

GENERAL INFORMATION:
APPLICANT: Pompeius, Markus
APPLICANT: Krieger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberer, Gregor
APPLICANT: Kim, Jun-Mon
APPLICANT: Lee, Heung-Schick
APPLICANT: Hwang, Byung-Joon
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
FILE REFERENCE: BGI-121CP2
CURRENT APPLICATION NUMBER: US/09/746,660A
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 09/606740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 09/603124
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/142101
PRIOR FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: 60/148613
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/187970
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
NUMBER OF SEQ ID NOS: 125
SOFTWARE: PatentIn Vers. 2.0
SEQ ID NO 97
LENGTH: 1557
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: CDS
LOCATION: (101)..(1534)
OTHER INFORMATION: RKN00132
US-09-746-660A-97

Query Match
Best Local Similarity 57.0%; Score 230.6; DB 9; Length 1557;
Matches 482; Conservative 0; Mismatches 354; Indels 9; Gaps 3;

QY 1011 GAAGAGAGGCTGACGTGCTTACAGGCTGTATCAGCTTCCAAAGCTGGGAAGCTCTGT 1070
DB 698 GAGGAAACACACACCGGTGTGACGCGCTGTACACCTCGCTGAAGAGCGCGCTGCGCT 757
QY 1071 GTCCGGCCATGAGCTCAATGATCTGTACCAAGAGAGTTGATTAAGTCTGACTGC 1130
DB 758 TTCACGCGATGACGTCAACGACGCTGTACACCAAGTTCATTAACAAGTACGGC 817
QY 1131 TCCCGAGATTCATTTGGATGGCTGAAGAGACACAGATGTGATTTGGTGGGAAA 1190
DB 818 ACCCGCACTCCCTGATGACGAGCAACCGCGCATCTGACATGCTCATGGCGGCAAG 877
QY 1191 CAAGT 1250
DB 878 AACGT 937
QY 1251 CTTGAGCAATGTGTACATTCAGAAATCGAACCCCATCTGTCTGTGAGGCGCTCATG 1310
DB 938 CAGGGCGCTCCGCTACAGGTACCGGAAGCTGACCCCATTAACGCTCTTGAAGCTGTGATG 997
QY 1311 GATGGTTTCAAGGTGTAAGCTTAATGAATCATCCGCAAGTGTGATGATTAAGT 1370
DB 998 GATGGCTACTGT 1057
QY 1371 TGCACAGGAATGAAGT 1430
DB 1058 GCGACCGGCAACAGACATCTATTTCTTCTGAGCAGATGCTCAAGATGAAGATCACGCT 1117
QY 1431 ATGTATGAATATGAGGCACTCCAAACAGAAATGATGTG---ACACGCTCCGCACT 1487

DB 1118 CTGCTGGGCAACATCGGTCTACCTTTGATTAATAGATCGATATGATTCCTCTGTGACCG 1177
QY 1488 CCGAGCTGAGGTGGAGAGGAGTACGTTCTCAGTGGACATGTCACTGGCCAGATGGC 1547
DB 1178 GACGACGTACCCGACACAGATCAAGCCACAGGTGACGAGTTTCACTTCCACCGGT 1237
QY 1548 AAGCAAGTGTCTCTGCGGACAGAGGTGTCTTACTCATTTAGCTGTCCAC---AGTT 1604
DB 1238 CGCTCATCATCTGCTGCTGCGAAGGTGCGCTGTGAACCTTGGCAAGCCACCGGACAC 1297
QY 1605 CCCACCTTTGTTCTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1664
DB 1298 CCATCATTTGTGATGTCAATCTTTGCGCCATGAGCAATGCGAGATGACATGCTGTC 1357
QY 1665 AATGACCCGAGGCGCATCAACAGAGATGTGCTACTGCTTCTCAAGAAATGATGATA 1724
DB 1358 CA---AAGCAGAGACAGTACAGAGAGAGGTGTACGCTGTGCTTGAAGTTCTGACGAA 1414
QY 1725 TACGTTGCCAGCTTGCATCTGCGATCATTTGATGATCCCACTTACAGAGCTGACAGATGAC 1784
DB 1415 AAGGTGACAGCATCTGAGGCTCTGCGGCTGAGTACAGCACTGACCAAGAG 1474
QY 1785 CAGCAAAATATCTGGGACATCAAAATATGAGGCTTCAACTTAATTTACAGATAC 1844
DB 1475 CAGCTGAGTACATGCGGCTGTGACGCTTGCAGGCCCATTCAGCGGAGACACTACGCTAC 1534
QY 1845 TAATG 1849
DB 1535 TAATG 1539

RESULT 12
US-09-738-626-836
Sequence 836, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: MARAGAMA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHINAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 836
LENGTH: 1422
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-836

Query Match
Best Local Similarity 56.8%; Score 225.6; DB 9; Length 1422;
Matches 477; Conservative 0; Mismatches 354; Indels 9; Gaps 3;

QY 1011 GAAGAGAGGCTGACGTGCTTACAGGCTGTATCAGCTTCCAAAGCTGGGAAGCTCTGT 1070
DB 586 GAGGAACACACACCGGTGTGACGCGCTGTACCACTTCCGTGAAGAGCGTGTCTCT 645

QY 1071 GTCCGGCCATGAACGTCATGATTCCTTACCAACAGAGTTGATTAATCTTACTGCG 1130
 DB 646 TTCCCAAGCCATGACGTCACGACGCTGTACCAAGTCACAGTTGATTAACAGTACGCG 705
 QY 1131 TCCCGAAGATTCATTTTGGATGCGCTGAAAGAGACACAGATGATGTTGGTGGAAA 1190
 DB 706 ACCCGGACACTCCCTGATGACGACATCAACCGGCGCACTGACATGCTCATGGGGGGAAG 765
 QY 1191 CAAGTGGT 1230
 DB 766 AACGCTCTGT 825
 QY 1251 CTGAGAGCAATGCTCATACATACGAAATGAGCCCATCTGTGTGTGTGTGTGTGTGTGTGT 1310
 DB 826 CAGGGGCGCTCGCTCAAGGTCACCGAAGCTGACCCATCAACGCTCTTCCAGCTGTGTGT 885
 QY 1311 GATGGTTCAGGGT 1370
 DB 886 GATGGTTCAGGGT 945
 QY 1371 TGCACAGGAATTAAGATGATGACACGCGGAGCATTGGATCGCATGAAAACAGTTGT 1430
 DB 946 GCGACCGGCAACAGGACATCATTTCTTCCAGAGATGCTCAAGATGAAAGATCAGCGCT 1005
 QY 1431 ATGCTATGCAATATGGGCGCTCCACACAGAAATGATGTG--ACGAGCCTCGGACT 1487
 DB 1006 CTGCTGGGCAACATCGCTCACTTTGATTAAGATGATGATGATGATGATGATGATGATG 1065
 QY 1488 CCGAGCTGACGTCGGAGGAGTACGTTCTCAGGTGAGCCATGTCATGTCGCGAGATGGC 1547
 DB 1066 GAGGAGGTGACCGGACACAGATCAACGCAAGGTGAGCGATGTCACCTTCCACCGGT 1125
 QY 1548 AAGAGTGTCTCTCTCTGCGAGAGGTGCTGCTACTCAATTTGAGCTGTCTCAG--AGTT 1604
 DB 1126 CGCTCCATCATCTCTCTCTCTGCGAGAGGTGCTGCTGTAACCTTGGCAACGCGGAGAC 1185
 QY 1605 CCCACCTTGTCTGTGTCATACACAGCCAAACAGGCTTGGACATGATGAATCTGAT 1664
 DB 1186 CCACTATTTGTCAATGTCACACTTTTCCGAGTACAGCATGCGCATGAGATGAATGATGTC 1245
 QY 1665 AATGACCCGAGGCGGAGTCAACAGAGATGATGATGATGATGATGATGATGATGATGATG 1724
 DB 1246 CA---AAGCAGAGAGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1302
 QY 1725 TACGTTGACAGTTCATGCTGACATCATTTGATGAGCCACCTTACAGAGCTGACAGATGAC 1784
 DB 1303 AAGGTGGCAGCAGTCCACGTTGAGGCTCTCGGCGGTGAGCTCACCGAATGACCAAGAG 1362
 QY 1785 CAGGCAAAATATCTGGGACTCAACAAAATGGGCGCATTCAACTATTTATTAAGATNC 1844
 DB 1363 CAGGCTGAGTACATCGCGCTTGAAGCTTGAAGGCCCATTCAGCGGAGACATCCGCTAC 1422

RESULT 13

US-10-021-121-1/c

Sequence 1, Application US/10021121

Patent No. US2002014244A1

GENERAL INFORMATION:

APPLICANT: Caras, Ingrid M

TITLE OF INVENTION: A2-1 Neurotrophic Factor

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/021.121
 FILING DATE: 06-Dec-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/635.130
 FILING DATE: 19-Mar-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Torchia, PhD., Timothy E.
 REGISTRATION NUMBER: 36,700
 REFERENCE/DOCKET NUMBER: P1001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650/225-8674
 TELEFAX: 650/952-9881
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1877 base pairs
 TYPE: Nucleic Acid
 STRANDEDNESS: Double
 TOPOLOGY: Linear
 FEATURE:
 NAME/KEY: Extra Cellular Domain
 LOCATION: 244-899
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 FEATURE:
 NAME/KEY: Transmembrane Domain
 LOCATION: 901-978
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 FEATURE:
 NAME/KEY: signal peptide
 LOCATION: 244-321
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-10-021-121-1
 Query Match 7.4%; Score 190; DB 12; Length 1877;
 Best Local Similarity 100.0%; Pred. No. 2.8e-43;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1283 CCCCATCTGCTGTCGCGAGGCGTCGATGAGGGGTTCAAGGGGTAAAGCTAAATGAGT 1342
 DB 1446 CCCCATCTGCTGTCGCGAGGCGTCGATGAGGGGTTCAAGGGGTAAAGCTAAATGAGT 1387
 QY 1343 CATCGGCAAGTCGATGCTGATTAATCTGACAGAAATAAGATGATGACAGCGGA 1402
 DB 1386 CATCGGCAAGTCGATGCTGATTAATCTGACAGAAATAAGATGATGACAGCGGA 1327
 QY 1403 GCACCTGGATGCGATGAATAACAGTTGATGCTATGCAATATGAGGCCACTCCACACAGA 1462
 DB 1326 GCACCTGGATGCGATGAATAACAGTTGATGCTATGCAATATGAGGCCACTCCACACAGA 1267
 QY 1463 AATGATGTG 1472
 DB 1266 AATGATGTG 1257

RESULT 14

US-09-746-660A-101

Sequence 101, Application US/09746660A

Publication No. US20030049804A1

GENERAL INFORMATION:

APPLICANT: Pompeius, Markus

APPLICANT: Schrodter, Hartwig

APPLICANT: Zelder, Oskar

APPLICANT: Haberhauser, Gregor

APPLICANT: Kim, Jun-Won

APPLICANT: Lee, Heung-Schick

APPLICANT: Hwang, Byung-Joon

TITLE OF INVENTION: CORINBACTERIUM GLUTAMICUM GENES ENCODING

GenCore version 5.1.4.P5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 14:00:04 ; Search time 3270 Seconds

(Without Alignment(s))
12693.894 Million cell updates/sec

Title: US-09-782-051-1

Perfect score: 2563
Sequence: 1 ggcgcggcgagctcgagctc.....aaaaaaaaaaaaaaaaaaaaa 2563

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

```
EST:
1: em_estha:*
2: em_esthm:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hnc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hnc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1321	51.5	1379	14	U69200 U69200 Soar
2	967.6	37.8	3234	11	AK014539 Mus muscu
3	933.4	36.4	1059	13	BM547143 AGENCOURT
4	926.2	36.1	938	9	AL527928 AL527928
5	897.4	35.0	921	9	AL551097 AL551097
6	896.2	35.0	957	9	AL520703 AL520703

Result No.	Score	Query Match	Length DB	ID	Description
7	884.4	34.5	950	9	AL527927 AL527927
8	873	34.1	919	9	AL520704 AL520704
9	868.6	33.9	917	9	AL577470 AL577470
10	868.4	33.9	916	9	AL574762 AL574762
11	862.6	33.7	1042	13	BM449470 AGENCOURT
12	856.2	33.4	865	9	AL519571 AL519571
13	847	33.0	967	13	BM468564 AGENCOURT
14	845.8	33.0	858	9	AL554850 AL554850
15	844.2	32.9	1021	14	BM059373 AGENCOURT
16	842.6	32.9	871	9	AL525362 AL525362
17	818.2	31.9	1042	13	BM452485 AGENCOURT
18	814	31.8	839	9	AL525407 AL525407
19	807.4	31.5	867	13	BM453920 AGENCOURT
20	800	31.2	861	9	AL519570 AL519570
21	785.8	30.7	839	14	BM212518 AGENCOURT
22	775.6	30.3	871	12	BM708016 AGENCOURT
23	775.4	30.3	833	9	AL510146 AGENCOURT
24	773.4	30.2	928	14	BM041487 AGENCOURT
25	762.2	29.7	857	13	BM684112 AGENCOURT
26	761.6	29.7	1036	13	BM466752 AGENCOURT
27	755.8	29.5	849	9	AL519576 AL519576
28	753.2	29.4	961	9	AL533969 AL533969
29	732.8	28.6	764	9	AL526038 AL526038
30	732.6	28.6	971	12	BM966580 AGENCOURT
31	732.2	28.6	894	14	BM086777 AGENCOURT
32	731.6	28.5	832	10	BM270773 AGENCOURT
33	730.4	28.5	832	10	BM270773 AGENCOURT
34	724.4	28.3	927	14	BM089473 AGENCOURT
35	723.4	28.2	843	13	BM457815 AGENCOURT
36	720.4	28.1	933	14	BM0889173 AGENCOURT
37	716.4	28.0	747	12	BM697210 AGENCOURT
38	713.2	27.8	744	12	BM699160 AGENCOURT
39	712.2	27.8	806	12	BM774719 AGENCOURT
40	706.2	27.6	741	13	BM333908 AGENCOURT
41	703.8	27.5	969	12	BM259448 AGENCOURT
42	703.2	27.4	937	14	BM0677104 AGENCOURT
43	703	27.4	1071	13	BM458493 AGENCOURT
44	702.4	27.4	801	13	BM335041 AGENCOURT
45	698.2	27.2	751	9	AL533711 AL533711

ALIGNMENTS

RESULT 1
U69200
LOCUS
DEFINITION U69200 Soares infant brain INIB Homo sapiens CDNA clone 23931, mRNA
ACCESSION U69200
VERSION U69200.1 GI:2739423
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (Bases 1 to 1379)
AUTHORS Volorio, S., Simon, G., Repetto, M., Cucchiardi, M., Banfi, S., Borsani, G., Ballabio, A. and Zollo, M.
TITLE Sequencing analysis of forty-eight human image cDNA clones similar to Drosophila mutant protein
JOURNAL DNA Seq. 9 (5-6), 307-315 (1998)
MEDLINE 99452388
COMMENT Contact: Zollo, Massimo
Telethon Institute of Genetics and Medicine
Via Olgettina 58, Milan, MI 20132, Italy
Email: zollo@tigem.it.

FEATURES

source
1. 1379
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1p13.3"
/clone="23931"

Query Match	Best Local Similarity	Matches 1349;	Conservative	0;	Mismatches	10;	Indels	2;	Gaps	2;
1201	51.5%	Score 1321;	DB 14;	Length 1379;						
1261	99.1%	Pred. No. 5.3e-229;								
3	0;									
1261	51.5%	Score 1321;	DB 14;	Length 1379;						
63	99.1%	Pred. No. 5.3e-229;								
1321	51.5%	Score 1321;	DB 14;	Length 1379;						
123	99.1%	Pred. No. 5.3e-229;								
1381	51.5%	Score 1321;	DB 14;	Length 1379;						
183	99.1%	Pred. No. 5.3e-229;								
1441	51.5%	Score 1321;	DB 14;	Length 1379;						
243	99.1%	Pred. No. 5.3e-229;								
1501	51.5%	Score 1321;	DB 14;	Length 1379;						
303	99.1%	Pred. No. 5.3e-229;								
1561	51.5%	Score 1321;	DB 14;	Length 1379;						
363	99.1%	Pred. No. 5.3e-229;								
1621	51.5%	Score 1321;	DB 14;	Length 1379;						
423	99.1%	Pred. No. 5.3e-229;								
1681	51.5%	Score 1321;	DB 14;	Length 1379;						
483	99.1%	Pred. No. 5.3e-229;								
1741	51.5%	Score 1321;	DB 14;	Length 1379;						
543	99.1%	Pred. No. 5.3e-229;								
1801	51.5%	Score 1321;	DB 14;	Length 1379;						
603	99.1%	Pred. No. 5.3e-229;								
1861	51.5%	Score 1321;	DB 14;	Length 1379;						
663	99.1%	Pred. No. 5.3e-229;								
1921	51.5%	Score 1321;	DB 14;	Length 1379;						
723	99.1%	Pred. No. 5.3e-229;								
1981	51.5%	Score 1321;	DB 14;	Length 1379;						

DB	783	TCATTATCCAAAGTTCGACAGACACAGAAAGTTCATTGGTGGCTTTAGATTAATA	842
QY	2041	GAAGTTGAGGGTCCCTCAGTCTAGTACTAAAGAAGATTTTACTCCCCAGCCAGAAA	2100
Db	843	GAAGTTGAGGGTCCCTCAGTCTAGTACTAAAGAAGATTTTACTCCCCAGCCAGAAA	902
QY	2101	GGTATTTCTCTCTTTACCATTTCGGGGGACTTTTGCTTAATTGAGTACCTTATTACA	2160
Db	903	GGTATTTCTCTCTTTACCATTTCGGGGGACTTTTGCTTAATTGAGTACCTTATTACA	961
QY	2161	GGAATGCTAAGGATACCTTCTCTGTGGAAACAATCGCAATGCTAAATCGCCTTAAAGA	2220
Db	962	GGAATGCTAAGGATACCTTCTCTGTGGAAACAATCGCAATGCTAAATCGCCTTAAAGA	1021
QY	2221	GCCATTTCTTAGTCTGTGAAATCAGTCTCTTTCACATTCTTCAGAGAAGCAGGATGGT	2280
Db	1022	GCCATTTCTTAGTCTGTGAAATCAGTCTCTTTCACATTCTTCAGAGAAGCAGGATGGT	1081
QY	2281	ACCTACCCGGCAGGATGATGATGATGGGTGGTGGATGTTAATTCCTTACAAATTCOA	2340
Db	1082	ACCTACCCGGCAGGATGATGATGATGGGTGGTGGATGTTAATTCCTTACAAATTCOA	1141
QY	2341	AGCCCTGTTTCCCTGGTAAAGGTGTATGTCCAGTTCAGAGATGTGTAAATAGACATGG	2400
Db	1142	AGCCCTGTTTCCCTGGTAAAGGTGTATGTCCAGTTCAGAGATGTGTAAATAGACATGG	1201
QY	2401	CTTGTAAAGATCAGAGAGGCCACTTGATTTATGATTAAGCCCTTCCTCACCACCA	2460
Db	1202	CTTGTAAAGATCAGAGAGGCCACTTGATTTATGATTAAGCCCTTCCTCACCACCA	1261
QY	2461	GACCTGTCATTTTTCGAGTTTAAACATGATACACCTATTTGAGTTATTTTGTCC	2520
Db	1262	GACCTGTCATTTTTCGAGTTTAAACATGATACACCTATTTGAGTTATTTTGTCC	1320
QY	2521	TCTAGATTATTTCTGTGTCCAAAAAATTTTTTTTTTAAAAA 2561	
Db	1321	TCTAGATTATTTCTGTGTCCAAAAAATTTTTTTTTTAAAAA 1361	
RESULT 2			
AK014539			
LOCUS			
DEFINITION	AK014539	3234 bp	mRNA linear HTC 19-JAN-2002
			Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched
			library, clone:4631427C17.homolog to PUTATIVE
			ADENOSYLHOMOCYSTEINASE (EC 3.5.1.1) (S-ADENOSYL-L-HOMOCYSTEINE
			HYDROLASE) (ADOMCYASE), full insert sequence.
ACCESSION	AK014539		
VERSION	AK014539.1	GI:12852459	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) 0 day neonate skin cDNA to mRNA,		
	clone:4631427C17.		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murine; Mus.		
REFERENCE			
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99379253		
PUBMED	10349636		
REFERENCE			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
TITLE	Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	11042159		
PUBMED	11042159		
REFERENCE			
AUTHORS	Shibata, K., Itoh, M., Alizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
TITLE	Kono, H., Akizawa, Y., Nishi, K., Katsuna, T., Teshiro, H., Itoh, M.,		

TITLE RIKEN integrated sequence analysis (RISA) system—384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

AUTHORS
Kawaji, Y., Shimagawa, A., Shibata, K., Yoshino, H., Itoch, M., Ishii, Y.,
Aikawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kijosawa, H., Kondo, S., Saito, R.,
Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R.,
Kado, K., Matsuda, H., Ashburner, M., Batailo, S., Casavan, T.,
Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiya, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nishido, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Stabill, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Botfield, D., Bojunga, N.,
Carrinci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamuya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzairelli, J., Kombarts, P., Nordone, P.,
Riing, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schobach, C., Seya, T., Shibata, S., Storch, K. F., Suzuki, H.,
Tayo-ok, K., Wang, K. H., Weitz, C., Whitaker, C., Wilmberg, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.
and Hayashizaki, Y.

TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409 (6821), 685-690 (2001)


```

OY 1737 TTGCATCTCCATCATTTATATGCCCCCTTACAGACATGACATGACCAAGCAAAATAT 1796
    |||||||
DB 380 TTGCATCTCCATCATTTATATGCCCCCTTACAGACATGACATGACCAAGCAAAATAT 439
OY 1797 CCGGGAGCTCAAAAAAATGGGCATTCACAAACCTAATATATACAGATACATAGGACATA 1856
    |||||||
DB 440 CCGGGAGCTCAAAAAAATGGGCATTCACAAACCTAATATATACAGATACATAGGACATA 499
OY 1857 CTACCAAGACAGACATCCACCTGACACACACACACCTTAAGAAATATTTTAAATAACTT 1916
    |||||||
DB 500 CTACCAAGACAGACATCCACCTGACACACACACACCTTAAGAAATATTTTAAATAACTT 559
OY 1917 TTAATTTCTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1976
    |||||||
DB 560 TTAATTTCTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 619
OY 1977 CATCTCATATCCATCATTTCTGACAGACACACAGAACTCTCTCTCTCTCTCTCTCTCT 2036
    |||||||
DB 620 CATCTCATATCCATCATTTCTGACAGACACACAGAACTCTCTCTCTCTCTCTCTCTCT 679
OY 2037 AATAGAGATTCAGAGGCTCCCTCACTAGTCACTAAAGAGATTTTCTCTCTCTCTCTCT 2096
    |||||||
DB 680 AATAGAGATTCAGAGGCTCCCTCACTAGTCACTAAAGAGATTTTCTCTCTCTCTCTCT 739
OY 2097 GAAAGGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2156
    |||||||
DB 740 GAAAGGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 798
OY 2157 AACAGAAATGCTAAAGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2216
    |||||||
DB 799 AACAGAAATGCTAAAGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 858
OY 2217 AAGAGCCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2276
    |||||||
DB 859 AAGAGCCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 918
OY 2277 TGTGACCTACCCCGCAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2334
    |||||||
DB 919 TGTGACCTACCCCGCAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 978
OY 2335 GTTCCAGAGCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2392
    |||||||
DB 979 GTTCCAGAGCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1038
OY 2393 GAGCATGCTGTGT 2405
    |||||||
DB 1039 GAGCATGCTGTGT 1051

```

```

RESULT 4
AL527928 938 bp mRNA linear EST 13-FEB-2001
LOCUS AL527928 LRI.NFL003.NBC3 Homo sapiens cDNA clone CSDDC027XC19 5
DEFINITION prime, mRNA sequence.
VERSION AL527928.1 GI:12791421
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 938)
AUTHORS Li, W.B., Gruber, C., Jessup, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
            Genoscope, Centre National de Sequencage
            BP 191 91006 Evry cedex - France
            Email: sequef@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES
            1..938
            /organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/clone="CSDDC027XC19"
/clone_lib="LRI.NFL003.NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by life technologies. Contact : Feng Liang life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 241 a 208 c 275 g 212 t 2 others
ORIGIN

```

```

Query Match 36.18; Score 926.2; DB 9; Length 938;
Best Local Similarity 99.7%; Pred. No. 1.3e-157;
Matches 936; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

OY 289 TCGGGGAGAGCTGAAGGCGCAAGAGATCGAGAGCGCCGAGAAAGTCTTCATNG 348
    |||||||
DB 1 TCGGGGAGAGCTGAAGGCGCAAGAGATCGAGAGCGCCGAGAAAGTCTTCATNG 60
OY 349 CCACCGTCACCAAGCGCCCAAGAGCAATCCAGTTGCTGATGACATGACAGAGTTCA 408
    |||||||
DB 61 CCACCGTCACCAAGCGCCCAAGAGCAATCCAGTTGCTGATGACATGACAGAGTTCA 120
OY 409 CCAATTCGCCCAAAATGCGGAGAGATTTTGTCTGCTCGATCTCACAGTCTCCA 468
    |||||||
DB 121 CCAATTCGCCCAAAATGCGGAGAGATTTTGTCTGCTCGATCTCACAGTCTCCA 180
OY 469 CTGACAGTACAGTTCAGCTCATCTACACAGATGCTGATGATGAGTTCTCTCC 528
    |||||||
DB 181 CTGACAGTACAGTTCAGCTCATCTACACAGATGCTGATGATGAGTTCTCTCC 240
OY 529 GAGAGAGCAACAACCAACTCCAAAGGCGACAGCAATTTCTGTGTAAGAAATCAAGC 588
    |||||||
DB 241 GAGAGAGCAACAACCAACTCCAAAGGCGACAGCAATTTCTGTGTAAGAAATCAAGC 300
OY 589 AGGCAGATTTGGACCGCGGAGATTTGATTTGCAAGACATGCTCTCTGATTT 648
    |||||||
DB 301 AGGCAGATTTGGACCGCGGAGATTTGATTTGCAAGACATGCTCTCTGATTT 360
OY 649 CACTCAGAAACGTGCTCAGGCGGAGAGAGCCCTGCTGCTGCTAAATAGTGGCTGA 708
    |||||||
DB 361 CACTCAGAAACGTGCTCAGGCGGAGAGAGCCCTGCTGCTGCTAAATAGTGGCTGA 420
OY 709 CACACATCACAGCCGACAGCGGCTGTTGATTGAGACACTGTGCCCTGGGGGCTCAGT 768
    |||||||
DB 421 CACACATCACAGCCGACAGCGGCTGTTGATTGAGACACTGTGCCCTGGGGGCTCAGT 480
OY 769 GCCGCTGCTGCTTGTAACTTACTCTAACTCAGAAATAGAGTACTCAGACCTGGCTG 828
    |||||||
DB 481 GCCGCTGCTGCTTGTAACTTACTCTAACTCAGAAATAGAGTACTCAGACCTGGCTG 540
OY 829 AGGCTGAGTGTGAGTGTCTGGAAGGCGAGTCTGGAAGTACTTGTGGTGTGA 888
    |||||||
DB 541 AGGCTGAGTGTGAGTGTCTGGAAGGCGAGTCTGGAAGTACTTGTGGTGTGA 600
OY 889 TTGACCGCTGTGTAACTGTGAGTGGTGGAGCCCAACATGATCTCTGATGATGGGGAG 948
    |||||||
DB 601 TTGACCGCTGTGTAACTGTGAGTGGTGGAGCCCAACATGATCTCTGATGATGGGGAG 660
OY 949 ACTTAACCATGGGTTTATAGAGTATCCAAAGCTTTTAAAGATCCGAGGATG 1008
    |||||||
DB 661 ACTTAACCATGGGTTTATAGAGTATCCAAAGCTTTTAAAGATCCGAGGATG 720
OY 1009 TGAAGAGAGGCTGATGCTTCAAGGCTGTATCAGCTCTCCAAAGCTGGAGGCTCT 1068
    |||||||

```

Dd	721	TGAGAGAGCGCTACTGGTGTTCACAGGCGCTATCAGCTCTCCAAAGTGGAAGCTCT	780
Oy	1069	GCTTCGCCGCCTGTAACGTCATATTCTGTGTACCACAAGAATTGATACTGTACT	1128
Dd	781	GGTTTCGCGCCATGACGTCATGATTCTGTATCCCAACAGAAAGTTGATACTGTACT	840
Oy	1129	GCTGCCGAATCCATTTTGGATGGCCCGAAGACCAACATGATGTTGGTGGGA	1188
Dd	841	GCTGCCGAATCCATTTTGGATGGCCCGAAGACCAACATGATGTTGGTGGG-	899
Oy	1189	AACAAGTGGTGTGTGTGCTATGGTGTGAGTAGGACAG	1227
Dd	900	AAMAAGTGGTGTGTGTGCTATGGTGTGAGTAGGACAG	938
RESULT 5			
LOCUS	AL551097	921 bp	mRNA linear EST 16-FEB-2001
DEFINITION	AL551097 LTR.NFL006.PL2 Homo sapiens cDNA clone CS0D1066TM07		5
ACCESSION	prime mRNA sequence.		
VERSION	AL551097		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 921) Li,M.B., Gruber,C., Jessee,J. and Polayes,D. Full-length cDNA libraries and normalization unpublished (2001)		
JOURNAL	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
COMMENT	Location/Qualifiers		
FEATURES			
SOURCE	1..921		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0D1066TM07"		
	/clone_1lb="LTR.NFL006.PL2"		
	/tissue_type="placenta"		
	/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com"		
BASE COUNT	217 a 231 c 302 g 170 t	1 others	
ORIGIN			
Query Match	35.0%; Score 897.4;	DB 9;	Length 921;
Best Local Similarity	99.6%;	Pred. No. 2,1e-152;	
Matches 909;	Conservative 1;	Mismatches 2;	Indels 1;
			Gaps 1;
Oy	69	GCCTTGGGCTGCCCAGAACAGACAGAGCGCTGGGCGCACAGCACTCAGAGCCGACGAGCTC	128
Dd	1	GCTTGGGCTGCCCAGAACAGACAGAGCGCTGGGCGCACAGCACTCAGAGCCGACGAGCTC	60
Oy	129	GACGAGGAGGCGCGAGAGAGGAGGTGGGCATGGCGTGTGAGAGGCGCGCGCGGCGCAAGC	188
Dd	61	GACGAGGAGGCGCGAGAGAGGAGGTGGGCATGGCGTGTGAGAGGCGCGCGCGGCGCAAGC	120
Oy	189	GAGGCGGCGCCAGAGAGGAGGAGGCGCGGCGCGGCGGTAGCCGCTGCGCGCGCGCGC	248
Dd	121	GAGGCGGCGCCAGAGAGGAGGAGGCGCGGCGCGGCGGTAGCCGCTGCGCGCGCGCGC	180
Oy	249	GGGGGAATGTGATGCTGACCGCATGCCGTGCCCGGGGTGCGGGAGAGAGCTGAAGCAG	308

Db	181	CGGGAGATTCGATGCTGACGCGATGCGGATGCGCCGGGCGTGGGGAGAGAGCTGAAAGCG	240
QY	309	GCCAGGAGATGAGAGACGCCGAGAAAGTACTCTTCATGTGGCCACCGTCAACAAAGCGCC	368
Db	241	GCCAGGAGATGAGAGACGCCGAGAAAGTACTCTTCATGTGGCCACCGTCAACAAAGCGCC	300
QY	369	AAGAAGCAATCCAGTTTGTGATGACATGACAGAGAGTTTACCAAAATTCGCCACCAAAAT	428
Db	301	AAGAAGCAATCCAGTTTGTGATGACATGACAGAGAGTTTACCAAAATTCGCCACCAAAAT	360
QY	429	GGCCGAGATCTTTGTCTGCTCGATGATCTCAGTCCCTCCACTGACAGCATGAGTTTACGT	488
Db	361	GGCCGAGATCTTTGTCTGCTCGATGATCTCAGTCCCTCCACTGACAGCATGAGTTTACGT	420
QY	489	GCACTCCACACAGATAGCTCTGATGATGATGAGGTTTCTCCCGAGAGAGCAGCAAAACAC	548
Db	421	GCACTCCACACAGATAGCTCTGATGATGATGAGGTTTCTCCCGAGAGAGCAGCAAAACAC	480
QY	549	TCCAAGGGGAGAGCAATTTCTGTGTGAGAAACATCAACACAGCAGAAATTTGGAGCGCG	608
Db	481	TCCAAGGGGAGAGCAATTTCTGTGTGAGAAACATCAACACAGCAGAAATTTGGAGCGCG	540
QY	609	GAGATTGAGATTGACAGACGACAAACATGTCGCTGTGATTTCACTCAGAGAAAGTGTGAC	668
Db	541	GAGATTGAGATTGACAGACGACAAACATGTCGCTGTGATTTCACTCAGAGAAAGTGTGAC	600
QY	669	GGGGAGAAACCCCTTGCTGCTGCTCTAAATATATGAGGGGTGTACACATCAACAGCCGAC	728
Db	601	GGGGAGAAACCCCTTGCTGCTGCTCTAAATATATGAGGGGTGTACACATCAACAGCCGAC	660
QY	729	GGGGGTGATTGAGACACTCTGTGCTGCTGGGGGCTCACTGCGCTGCTGCTGTGTAC	788
Db	661	GGGGGTGATTGAGACACTCTGTGCTGCTGGGGGCTCACTGCGCTGCTGCTGTGTAC	720
QY	789	ATCTACTCACTCAGAAATGATAGTGTGAGACTGCGAGACTGCGAGAGCTGGAATTCAGATTC	848
Db	721	ATCTACTCACTCAGAAATGATAGTGTGAGACTGCGAGACTGCGAGAGCTGGAATTCAGATTC	780
QY	849	GCTTGGAAAGGGGCAATCAGAAATGATCTCTGCTGTGTATGACCGCTGTGTGAACATG	908
Db	781	GCTTGGAAAGGGGCAATCAGAAATGATCTCTGCTGTGTATGACCGCTGTGTGAACATG	840
QY	909	GATGGGTGGCAGCGCAACATGATCTGTGATGATGAGGGGAGACTTAACCCACTGGTTAT	968
Db	841	GATGGGTGGCAGCGCAACATGATCTGTGATGATGAGGGGAGACTTAACCCACTGGTTAT	899
QY	969	AAGAGTATCCAA 981	
Db	900	AAGAAGTATCCAA 912	
RESULT 6			
AL520703/c	AL520703	957 bp	linear
LOCUS	AL520703	LTI_NFL004	NBC2 Homo sapiens cDNA clone CSDB002YL22 3
DEFINITION	prime, mRNA sequence.		
ACCESSION	AL520703		
VERSION	AL520703.1	GI:12784196	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo;		
TITLE	1 (bases 1 to 957)		
JOURNAL	L.I.W.B., Gruber, C., Jeejee, J. and Polayes, D.		
COMMENT	Full-length cDNA libraries and normalization		
	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 Evry cedex - France		
	Email: seque@genoscope.cns.fr, web : www.genoscope.cns.fr		
FEATURES	Location/Qualifiers		
source	1..957		


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0B0027C19"
/clone_1lb="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by life technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8371 Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com"
BASE COUNT      299 a      176 c      227 g      254 t      1 others
ORIGIN
Query Match      35.0%; Score 896.2; DB 9; Length 957;
Best Local Similarity 99.1%; Pred. No. 3.4e-152;
Matches 942; Conservative 1; Mismatches 4; Indels 4; Gaps 4;
OY 1572 GGTGCTACTCAATTTGAGCGTCCACAGTCCACCTTTGTCGTCAACAGCC 1631
|||||
DB 947 GGTGCTACTCAATTTGAGCGTCCACAGTCCACCTTTGTCGTCAACAGCC 889
OY 1632 ACAACACAGCGTTGGCACTGATAGACTCTAATATGCAACCGAGGGCGATACAGCAG 1691
|||||
DB 888 ACAACACAGCGTTGGCACTGATAGACTCTAATATGCAACCGAGGGCGATACAGCAG 829
OY 1692 GATGTGACTTGTCTTCCATAAGAAATGATGATAGTGGCCAGCTGTCATCGCATCA 1751
|||||
DB 828 GATGTGACTTGTCTTCCATAAGAAATGATGATAGTGGCCAGCTGTCATCGCATCA 769
OY 1752 TTTGATGCCACCTTACAGAGTGCAGATGACCAAGCAAAATATCTGGAGCTCAACAA 1811
|||||
DB 768 TTTGATGCCACCTTACAGAGTGCAGATGACCAAGCAAAATATCTGGAGCTCAACAA 709
OY 1812 AATGGGCAATTCATTAATTTATTAAGATTAATGAGCAATACATACCAAGAGCAGT 1871
|||||
DB 708 AATGGGCAATTCATTAATTTATTAAGATTAATGAGCAATACATACCAAGAGCAGT 649
OY 1872 CCACTGACACACACACTTAAGAAATTTTAAAGATTAATCTTATCTTCTTAC 1931
|||||
DB 648 CCACTGACACACACACTTAAGAAATTTTAAAGATTAATCTTATTTCTTCTTAC 589
OY 1932 TCCCTTCTCTGATTTTTCCTAATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCA 1991
|||||
DB 588 TCCCTTCTCTGATTTTTCCTAATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCA 529
OY 1992 GTTTCGACACACACAGAACTGCTTCATGCTCTTATGATGAATGAAGTTGAGG 2051
|||||
DB 528 GTTTCGACACACACAGAACTGCTTCATGCTCTTATGATGAATGAAGTTGAGG 469
OY 2052 TCCCTGACTTACTTACTTAAAGAGATTTTACTCCCGACCCGAAAGGTGATCTTC 2111
|||||
DB 468 TCCCTGACTTACTTACTTAAAGAGATTTTACTCCCGACCCGAAAGGTGATCTTC 410
OY 2112 TCTTACCATTTCTGGGACTTATGCTAATAGTACCTTATTAACAGAAATCTTA 2171
|||||
DB 409 TCTTACCATTTCTGGGACTTATGCTAATAGTACCTTATTAACAGAAATCTTA 350
OY 2172 GGTACCTTCTCTGTGAACAATCTCAATGTCTAATATGCTTAAAGAGCCATTTCTT 2231
|||||
DB 349 GGTACCTTCTCTGTGAACAATCTCAATGTCTAATATGCTTAAAGAGCCATTTCTT 290
OY 2232 AGCTGCTGAATCAATGCTCTTCTTCTTCAAGAGAGGAGATGATACCTACCCGCG 2291
|||||
DB 289 AGCTGCTGAATCAATGCTCTTCTTCTTCAAGAGAGGAGATGATACCTACCCGCG 230
OY 2292 AGGTAGGTAGATGTGGGTGTGCAATGTAATTTCCCTTAAGAGTTCAAGCCTGTTC 2351

```

```

|||||
DB 229 AGGTAGGTAGATGTGGGTGTGCAATGTAATTTCCCTTAAGAGTTCAAGCCTGTTC 170
OY 2352 GTGGTAAGGTGATGTGTCAGATTCAGAGATGTGATATGAGATGGCTTTAAGAT 2411
|||||
DB 169 GTGGTAAGGTGATGTGTCAGATTCAGAGATGTGATATGAGATGGCTTTAAGAT 110
OY 2412 CAGGAGCCCACTGATTTATATAGTATAGCCCTTCCACACGCCAGACTGTCAT 2471
|||||
DB 109 CAGGAGCCCACTGATTTATATAGTATAGCCCTTCCACACGCCAGACTGTCAT 50
OY 2472 TTTTCAGTTTATATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 2522
|||||
DB 49 TTTTCAGT-TTTTACTAGTACTACTCTA-TTGGAGTTATTTTGTCTC 1
RESULT 7
AL527927/c 950 bp mRNA linear EST 13-FEB-2001
LOCUS AL527927 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC027YC19 3
DEFINITION prime, mRNA sequence.
ACCESSION AL527927
VERSION AL527927.1 GI:12791420
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 950)
AUTHORS Li, W.-B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seque@genoscope.cns.fr, Web: www.genoscope.cns.fr.
FEATURES
source
1..950
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC027YC19"
/clone_1lb="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by life technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8371 Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com"
BASE COUNT      291 a      177 c      223 g      255 t      4 others
ORIGIN
Query Match      34.5%; Score 884.4; DB 9; Length 950;
Best Local Similarity 99.1%; Pred. No. 4.6e-150;
Matches 907; Conservative 2; Mismatches 4; Indels 2; Gaps 2;
OY 1572 GGTGCTACTCAATTTGAGCGTCCACAGTCCACCTTTGTCGTCAACAGCC 1631
|||||
DB 913 GGTGCTACTCAATTTGAGCGTCCACAGTCCACCTTTGTCGTCAACAGCC 855
OY 1632 ACAACACAGCGTTGGCACTGATAGACTCTAATATGCAACCGAGGGCGATACAGCAG 1691
|||||
DB 854 ACAACACAGCGTTGGCACTGATAGACTCTAATATGCAACCGAGGGCGATACAGCAG 795
OY 1692 GATGTGACTTGTCTTCCATAAGAAATGATGATAGTGGCCAGCTGTCATCGCATCA 1751
|||||
DB 794 GATGTGACTTGTCTTCCATAAGAAATGATGATAGTGGCCAGCTGTCATCGCATCA 735

```

QY 1752 TTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATGTGGACTCAACAA 1811
 DB 734 TTGTATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATGTGGACTCAACAA 675
 QY 1812 AATGGGCAATTAACCTATATTATACATATATGACCACTACTACCAAGACAGC 1871
 DB 674 AATGGGCAATTAACCTATATTATACATATATGACCACTACTACCAAGACAGC 615
 QY 1872 CCACCTGAACACACACACTTAAGAAATATTTTATAGATACTTTATTTCTCTTAC 1931
 DB 614 CCACCTGAACACACACACTTAAGAAATATTTTATAGATACTTTATTTCTCTTAC 555
 QY 1932 TCCCTTCTCTGATTTTCTCTATATTTTCAATCTCTTTTTCATCTCATTTACCA 1991
 DB 554 TCCCTTCTCTGATTTTCTCTATATTTTCAATCTCTTTTTCATCTCATTTACCA 495
 QY 1992 GTTTCGACACACACAGAACTTCTTATGCTTCTTATAGAAATAGAAATTCAGG 2051
 DB 494 GTTTCGACACACACAGAACTTCTTATGCTTCTTATAGAAATAGAAATTCAGG 435
 QY 2052 TCCCTCACTCTAGTCACTAAGAGATTTTACTCCCAAGCCCAAGAGATCTTTC 2111
 DB 434 TCCCTCACTCTAGTCACTAAGAGATTTTACTCTCCCAAGCCCAAGAGATCTTTC 376
 QY 2112 TCTTACCACTTCTGAGGACTTATGCTTATAGTACCTTATTAACAGAAATGCTAA 2171
 DB 376 TCTTACCACTTCTGAGGACTTATGCTTATAGTACCTTATTAACAGAAATGCTAA 316
 QY 2172 GTTACTCTCTGAGGAAATCTGCAATCTGCAATCTGCAATCTGCAATCTGCAAT 2231
 DB 316 GTTACTCTCTGAGGAAATCTGCAATCTGCAATCTGCAATCTGCAATCTGCAAT 256
 QY 2232 ACTCTGTAATCAGTCTCTTCACTCTTCAAGAGAGAGATGATCTTACCCGCG 2291
 DB 256 ACTCTGTAATCAGTCTCTTCACTCTTCAAGAGAGAGATGATCTTACCCGCG 196
 QY 2292 AGGTAGTATGATGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 2351
 DB 196 AGGTAGTATGATGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 136
 QY 2352 CAGGCGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2411
 DB 136 CAGGCGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 76
 QY 2412 CAGGAGGCCACTGATGATGATGATGATGATGATGATGATGATGATGATG 2471
 DB 76 CAGGAGGCCACTGATGATGATGATGATGATGATGATGATGATGATGATG 16
 QY 2472 TTTTGGAGTTTTAA 2486
 DB 15 TTTTGGAGTTTTAA 1
 RESULT 8
 AL520704 891 bp mRNA linear EST 13-FEB-2001
 LOCUS AL520704 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB002YL22 5
 DEFINITION Prime, mRNA sequence.
 ACCESSION AL520704
 VERSION AL520704.1 GI:12784197
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 891)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France

FEATURES
 source
 Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
 Location/Qualifiers
 1..891
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CS0DB002YL22"
 /clone_lib="LTI_NFL004_NBC2"
 /sex="male"
 /issue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen, 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@life.com
 http://fulllength.invitrogen.com

BASE COUNT 194 a 238 c 299 g 158 t 2 others
 ORIGIN

Query Match 34.1%; Score 873; DB 9; Length 891;
 Best Local Similarity 99.5%; Pred. No. 5,4e-148;
 Matches 873; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCGGGGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAG 60
 DB 15 GGGCGGGGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAG 74
 QY 61 GTCCGGTCTTGGGCTGCGCAACAGACAGAGGCGGCGACAGACCTCAGAGCGGA 120
 DB 75 GTCCGGTCTTGGGCTGCGCAACAGACAGAGGCGGCGACAGACCTCAGAGCGGA 134
 QY 121 CGCAGCTCGACAGAGGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
 DB 135 CGCAGCTCGACAGAGGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 194
 QY 181 GGGCAGGCGGGGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 DB 195 GGGCAGGCGGGGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 254
 QY 241 GGGCAGGCGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 DB 255 GGGCAGGCGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 314
 QY 301 TGAAGCAGGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 DB 315 TGAAGCAGGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 374
 QY 361 AGCGCGCCAGAGAGCAATTCAGTTGCTGATGATGATGATGATGATGATGATGATG 420
 DB 375 AGCGCGCCAGAGAGCAATTCAGTTGCTGATGATGATGATGATGATGATGATGATG 434
 QY 421 CCAAAATGCGGAGATCTTTGTCTGCTGATCTCAAGTCCCTCACTGACAGTACA 480
 DB 435 CCAAAATGCGGAGATCTTTGTCTGCTGATCTCAAGTCCCTCACTGACAGTACA 494
 QY 481 GTTCAGTGCATCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 DB 495 GTTCAGTGCATCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 554
 QY 541 AAACCACTCCAGAGGAGAGCAATTTCTGTGAAGAACATCAAGAGAGAGAGAGAG 600
 DB 555 AAACCACTCCAGAGGAGAGCAATTTCTGTGAAGAACATCAAGAGAGAGAGAGAG 614
 QY 601 GACCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 DB 615 GACCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 674
 QY 661 GTCTCAAGGGGAGAGCCCTTGGCTGCTTAAAGAGGAGGAGGAGGAGGAGGAGGAG 720

```

Db 675 GTGTCAGAGGGAGAGAGCCCTTGCTGCTAAATAGTGGCTGTACACACATCAG 734
QY 721 CCCAGACAGCCGCTGTATTTAGACACCTGTGCCCTGGGGCTACATGCCCTGGTCTG 780
Db 735 CCCAGACAGCCGCTGTATTTAGACACCTGTGCCCTGGGGCTACATGCCCTGGTCTG 794
QY 781 CTTGTAACATCTACCACTCAGATGAAAGTGTGACAGCACTGGCTGAGGTGAGTTG 840
Db 795 CTTGTAACATCTACCACTCAGATGAAAGTGTGACAGCACTGGCTGAGGTGAGTTG 854
QY 841 CAGTGTGCTTGGAGAGGCGAGTCAGAGATGACTT 877
Db 855 CAGTGTGCTTGGAGAGGCGAGTCAGAGATGACTT 891

RESULT 9
AL577470/c 917 bp mRNA linear EST 16-FEB-2001
LOCUS AL577470 LTI.NFL006.PL2 Homo sapiens cDNA clone CSOD1087YL18 3
DEFINITION prime, mRNA sequence.
ACCESSION AL577470
VERSION AL577470.1 GI:12940631
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 917)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Source
1. 917
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1id="LTI.NFL006.PL2"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@life.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 280 a 171 c 210 g 247 t 9 others
ORIGIN
Query Match 33.9%; Score 868.6; DB 9; Length 917;
Best Local Similarity 97.3%; Pred. NO. 3.4e-147;
Matches 885; Conservative 7; Mismatches 17; Indels 1; Gaps 1;

QY 1594 GCTCCAGAGTCCACCTTGTCTGTCATCAGACCAACAGAGGGCTTGGCACTGA 1653
Db 917 GCTCCAGAGTCCACCTTGTCTGTCATCAGACCAACAGAGGGCTTGGCACTGA 888
QY 1654 TAGAAGCTATATGACACCCGAGGGGCGATAGAGAGAGATGTACTTGTCTCTAAGA 1713
Db 857 TAGAAGCTATATGACACCCGAGGGGCGATAGAGAGAGATGTACTTGTCTCTAAGA 798
QY 1714 AAATGATGATATGCTGGCAGCTTGCATCTGCATCATTTGATGCCACCTTACAGAGC 1773
Db 797 AAATGATGATATGCTGGCAGCTTGCATCTGCATCATTTGATGCCACCTTACAGAGC 728
QY 1774 TGACGATGACCAAGCAAAATATCTGGGAGCTCAACAAAATGGCCATTCAACCTAAT 1833
Db 1833 TGACGATGACCAAGCAAAATATCTGGGAGCTCAACAAAATGGCCATTCAACCTAAT

```

```

Db 737 TSACAGATGACCAAGCAAAATATCTGGGAGCTCAACAAAATGGCCATTCAACCTAAT 678
QY 1834 ATTACAGATGATATGAGACATCTACCAAGAGCCAGTCCAGCTGAACCAACACTTA 1893
Db 677 ATTACAGATGATATGAGACATCTACCAAGAGCCAGTCCAGCTGAACCAACACTTA 618
QY 1894 AGAAATATTTTAAATTAATCTTTATTTTCTTACCTTACCTTCTGATTTTTC 1953
Db 617 AGAAATATTTTAAATTAATCTTTATTTTCTTACCTTACCTTCTGATTTTTC 558
QY 1954 CATATATTTCAATCTGTTTTCATCTCATATACCAAGTTGACAGACACACAGAGAC 2013
Db 557 CATATATTTCAATCTGTTTTCATCTCATATACCAAGTTGACAGACACACAGAGAC 498
QY 2014 TTGCTTCATAGGCTTTTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 2073
Db 497 TTGCTTCATAGGCTTTTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 438
QY 2074 AAGGATTTTACTCCCGAGCCAGAAAGGATTTCTCTTACATTTCTGGGAGCT 2133
Db 437 AAGGATTTTACTCCCGAGCCAGAAAGGATTTCTCTTACATTTCTGGGAGCT 379
QY 2134 TAGCTTTAATTAAGTACTTATTAACAGAAATGCTAAGTACCTCTCTGGAGCAAT 2193
Db 378 TAGCTTTAATTAAGTACTTATTAACAGAAATGCTAAGTACCTCTCTGGAGCAAT 319
QY 2194 CTGCATGCTCAATATGCGCTTAAAGAGCCCATTTCTTACTGCTGTAACAGTCTT 2253
Db 318 CTGCATGCTCAATATGCGCTTAAAGAGCCCATTTCTTACTGCTGTAACAGTCTT 259
QY 2254 TCACCTCTTCAAGAGAGAGAGAGTGTACCTACCGGAGAGTGTAGATGTGGTGT 2313
Db 258 TCACCTCTTCAAGAGAGAGAGAGTGTACCTACCGGAGAGTGTAGATGTGGTGT 199
QY 2314 GCATGTTAATTTCCCTTAGAAGTCCAGAGCTGTTCTGCGTAAGGTGTATGCCA 2373
Db 198 GCATGTTAATTTCCCTTAGAAGTCCAGAGCTGTTCTGCGTAAGGTGTATGCCA 139
QY 2374 GTTCAGAGAGTGTATTAATAGAGAGCTGTTAAGATGAGAGAGGCGCACTGATTTAT 2433
Db 138 GTTCAGAGAGTGTATTAATAGAGAGCTGTTAAGATGAGAGAGGCGCACTGATTTAT 79
QY 2434 AGTATAGCCCTTCTCCACCTCCACAGAGCTGCTATTTTTCAGATTTTAACTAGCT 2493
Db 78 AGTATAGCCCTTCTCCACCTCCACAGAGCTGCTATTTTTCAGATTTTAACTAGCT 19
QY 2494 ACACTCTATT 2503
Db 18 ACACTCTATT 9

RESULT 10
AL574762/c 916 bp mRNA linear EST 16-FEB-2001
LOCUS AL574762 LTI.NFL006.PL2 Homo sapiens cDNA clone CSOD1066YM07 3
DEFINITION prime, mRNA sequence.
ACCESSION AL574762
VERSION AL574762.1 GI:12935277
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 916)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Source
1. 916

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSOD1066Y07"
/clone_lib="LT1_NFL006_P12"
/tissue_type="Placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
Email: filangellif@tech.com URL:
http://fulllength.invitrogen.com"

BASE COUNT      267 a      174 c      215 g      246 t      14 others
ORIGIN
Query Match      33.9%; Score 868.4; DB 9; Length 916;
Best Local Similarity 97.9%; Pred. No. 3.6e-147;
Matches 899; Conservative 11; Mismatches 5; Indels 3; Gaps 3;

QY 1536 TGGCCAGATGCGAAGAGAGTGTCTCTCTGCGAGAGGTCGTCTACTGATTTGAGCTGC 1595
DB 916 TGGCCAGATGCGAAGAGAGTGTCTCTCTGCGAGAGGTCGTCTACTGATTTGAGCTGC 858
QY 1596 TCCACAGTCCCACTTTGTCTGTCTCATCAGCCCAACACAGAGCTTTGGCACTGATA 1555
DB 857 TCCACAGTCCCACTTTGTCTGTCTCATCAGCCCAACACAGAGCTTTGGCACTGATA 798
QY 1656 GAACCTATATATGACACCCGAGGGGCGATACACAGAGATGTGATCTGCTCTTAAGAAA 1715
DB 797 GAACCTATATATGACACCCGAGGGGCGATACACAGAGATGTGATCTGCTCTTAAGAAA 738
QY 1716 ATGATGATATACGTTGCGAGCTGTCATCTGATCTGATTTGATGATCCCACTTACAGAGCTG 1775
DB 737 ATGATGATATACGTTGCGAGCTGTCATCTGATCTGATTTGATGATCCCACTTACAGAGCTG 678
QY 1776 ACAGATGACCAAGCAAAATCTGGGACCTCAACAAAATGGGGCATTCACAACTATAT 1835
DB 677 ACAGATGACCAAGCAAAATCTGGGACCTCAACAAAATGGGGCATTCACAACTATAT 618
QY 1836 TACAGATATATGACATACCTACCAAGAGAGTGCACCTGACACACACTTAAG 1895
DB 617 TACAGATATATGACATACCTACCAAGAGAGTGCACCTGACACACACTTAAG 558
QY 1896 AAATATTTTAAAGATACCTTTATTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 1955
DB 557 AAATATTTTAAAGATACCTTTATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 498
QY 1956 AATAATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 2015
DB 497 AATAATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 438
QY 2016 GCTTCATGCTCTTTAGATGAATAGAGTTCAGGGTCCCTCACTAGTACATAAGAA 2075
DB 437 GCTTCATGCTCTTTAGATGAATAGAGTTCAGGGTCCCTCACTAGTACATAAGAA 378
QY 2076 GGATTTTACTCCCGACGCGCAAGAAAGTATCTTCTTTACCACTTTCTGGGAGCTTTA 2135
DB 377 GGATTTTACTCCCGACGCGCAAGAAAGTATCTTCTTTACCACTTTCTGGGAGCTTTA 319
QY 2136 GCTTAAATTAAGTACTTTTAAAGAGAAATGCTAAGTACCTCTCTGTGAACAATCT 2195
DB 318 GCTTAAATTAAGTACTTTTAAAGAGAAATGCTAAGTACCTCTCTGTGAACAATCT 259
QY 2196 GCATATGCTAATGCGCTTAAAGAGCCATTTCTTAGTGTGGAATGAGTCTCTTTC 2255
DB 258 GCATATGCTAATGCGCTTAAAGAGCCATTTCTTAGTGTGGAATGAGTCTCTTTC 199
QY 2256 ACTTCTTCAAGAGACAGGATGTAACCTACCGGAGGTAGGTATGATGTGGGTGTG 2315
DB 198 ACTTCTTCAAGAGACAGGATGTAACCTACCGGAGGTAGGTATGATGTGGGTGTG 139

```

```

QY 2216 ATGTATATTTCCCTTAAGATTCACAGCCCTTTCTTCTGCTTAAGGTGATGTCCAGT 2375
DB 138 ATGTATATTTCCCTTAAGATTCACAGCCCTTTCTTCTGCTTAAGGTGATGTCCAGT 79
QY 2376 TCGAGATGTGTATTAAGAGATGAGTGTGTAATATCAGCA-GGCCCACTGAGATTATA 2434
DB 78 TCGAGATGTGTATTAAGAGATGAGTGTGTAATATCAGCA-GGCCCACTGAGATTATA 19
QY 2435 GATATGCTCTCTCTCCAC 2452
DB 18 GATATGCTCTCTCCAC 1

RESULT 11
LOCUS      BM449470      1042 bp      mRNA      linear      EST 05-FEB-2002
DEFINITION AGENCOURT_6400881 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5493677
ACCESSION  BM449470
VERSION     BM449470.1 GI:18498510
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 1042)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.lnl.gov
              Plate: LLM12117 row: m column: 06
              High quality sequence stop: 694.
              Location/Qualifiers
                1..1042
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5493677"
                  /clone_lib="NIH_MGC_67"
                  /tissue_type="retinoblastoma"
                  /lab_host="DH10B (phage-resistant)"
                  /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                  Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                  Average insert size 1.75 kb. Library constructed by Life
                  Technologies."

BASE COUNT      267 a      219 c      309 g      240 t      7 others
ORIGIN
Query Match      33.7%; Score 862.6; DB 13; Length 1042;
Best Local Similarity 96.9%; Pred. No. 3.9e-146;
Matches 908; Conservative 0; Mismatches 25; Indels 4; Gaps 3;

QY 532 AGAAGCAGCAACCACTCCAGGCGAGCAGCAATTTCTGTGTAACAATCAAGAGG 591
DB 1 AGAAGCAGCAACCACTCCAGGCGAGCAGCAATTTCTGTGTAACAATCAAGAGG 60
QY 592 CAGAAATTTGAGCCCGGAGATGAGATTGACAGCAAGACATGCTGCTGATTTTAC 651
DB 61 CAGAAATTTGAGCCCGGAGATGAGATTGACAGCAAGACATGCTGCTGATTTTAC 120
QY 652 TCGAAGAACTGCTCAGGGGAGAGCCCTTGGCTGTCTAAATAGTGGCTGTACAC 711
DB 121 TCGAAGAACTGCTCAGGGGAGAGCCCTTGGCTGTCTAAATAGTGGCTGTACAC 180
QY 712 ACATCAAGCCAGACAGCGGTGTGATTCAGACACCTGTGCCCTGGGGCTCAGTGC 771

```

```

Db 181 ACATCACAGCCAGACACGGGTGTGATGACACTCTGTCCCTGGGGCTCAGTCC 240
OY 772 GCTGTCGCTTGTACATCTACTCACTCAGAAATGAGTACGACACTGGCTGAG 831
Db 241 GCTGTCGCTTGTACATCTACTCACTCAGAAATGAGTACGACACTGGCTGAG 300
OY 832 CTGAGTGCAGATGTCCTTGGAAAGGCGAGTCAAGAGTACTCTCTGGTGTATG 891
Db 301 CTGAGTGCAGATGTCCTTGGAAAGGCGAGTCAAGAGTACTCTCTGGTGTATG 360
OY 892 ACCGCTGTGACATGAGTGGGTGGAGGCCACATGATCTGTGATGAGGGAGACT 951
Db 361 ACCGCTGTGACATGAGTGGGTGGAGGCCACATGATCTGTGATGAGGGAGACT 420
OY 952 TAACCCATGGGTTATAGAGATCCAAAGCTTTAAGAGATCCAGAGCATTTGTG 1011
Db 421 TAACCCATGGGTTATAGAGATCCAAAGCTTTAAGAGATCCAGAGCATTTGTG 480
OY 1012 AAGAGAGCTGATGCTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGCTGTG 1071
Db 481 AAGAGAGCTGATGCTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGCTGTG 540
OY 1072 TTCCGGCCATGACGTCATGATCTGTACCAACAGAGTTGATTAATCTGTACTGT 1131
Db 541 TTCCGGCCATGACGTCATGATCTGTACCAACAGAGTTGATTAATCTGTACTGT 600
OY 1132 GCCAGATCCATTTTGATGAGGCTGGAAGAGACACAGATGTGATTTGGTGGAAAC 1191
Db 601 GCCAGATCCATTTTGATGAGGCTGGAAGAGACACAGATGTGATTTGGTGGAAAC 660
OY 1192 AAGTGTGTGTGTGATGAGTGTGAGTGTGAGGAGGCTGTGTCTGTCAAGCTC 1251
Db 661 AAGTGTGTGTGTGATGAGTGTGAGTGTGAGGAGGCTGTGTCTGTCAAGCTC 720
OY 1252 TTGAGCAATGTCATACATACCGAATCGAACCCATCTGTCTGTGAGGCTGTGATG 1311
Db 721 TTGAGCAATGTCATACATACCGAATCGAACCCATCTGTCTGTGAGGCTGTGATG 780
OY 1312 ATGGG--TTGAGGTCGTGTAAGCTTAATGAGTCAATCGGCAAGTCGATGTGTAATAC 1369
Db 781 ATGGGTCGTGTAAGCTTAATGAGTCAATCGGCAAGTCGATGTGTAATAC 840
OY 1370 TTGACAGGAATTAAGATGATGACACGAGGAGCACTT--GGATGGCATGAAAAACAGT 1428
Db 841 TTGACAGGAATTAAGATGATGACACGAGGAGCACTTGGATGGCATGAAAAACAGT 900
OY 1429 GTATC--GTATCAATATGGCCACTCCAAACAGAAA 1464
Db 901 GTATCAGTATGCAATATGGGGCAGCTCAACACAGAAA 937

RESULT 12
AL519571 865 bp mRNA linear EST 13-FEB-2001
LOCUS AL519571 LTI_NF004_NBC2 Homo sapiens cDNA clone CS0DB004Y104 5
DEFINITION prime, mRNA sequence.
ACCESSION AL519571
VERSION AL519571.1 GI:12783064
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 865)
AUTHORS Li, M. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers

```

```

source 1. 865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB004Y104"
/clone_1lb="LTI_NF004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="PH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 193 a 231 c 155 t 2 others
ORIGIN
Query Match 33.4%; Score 856.2; DB 9; Length 865;
Best Local Similarity 99.2%; Pred. No. 6e-145;
Matches 858; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
OY 8 GCAGTGGAGCTCGGAGCTGCTTCTGTCTGTGGCCACCGTGTCCGGC 67
Db 1 GCAGTGGAGCTCGGAGCTGCTTCTGTCTGTGGCCACCGTGTCCGGC 60
OY 68 TGCCTTGGGTGCGCAACAGACAGAGCGGTGGCCACAGCACTCAGAAAGCCGACGT 127
Db 61 TGCCTTGGGTGCGCAACAGAGCGGTGGCCACAGCACTCAGAAAGCCGACGT 120
OY 128 CGACGAGGGGCCGCGAGAGGAGGTGGCGCATCCGTGTGCGAGGGCGCCGCGGGCAGG 187
Db 121 CGACGAGGGGCCGCGAGAGGAGGTGGCGCATCCGTGTGCGAGGGCGCCGCGGGCAGG 180
OY 188 CGGGCGGGGCCGAGAGGGGAAAGAGCGGGCGGGGTACGCGCTGGCGGGCGGG 247
Db 181 CGGGCGGGGCCGAGAGGGGAAAGAGCGGGCGGGGTACGCGCTGGCGGGCGGG 240
OY 248 CGGGGAAATGTCGATCCCTGACGCGATGCGCTGCCGCGGGGTGGGGAGAGCTGAAGCA 307
Db 241 CGGGGAAATGTCGATCCCTGACGCGATGCGCTGCCGCGGGGTGGGGAGAGCTGAAGCA 300
OY 308 GGCAGAGAGATGAGAGCGCGGAGAGATCTCTTCATGCGCAGCTCAACAGCGGCG 367
Db 301 GGCAGAGAGATGAGAGCGCGGAGAGATCTCTTCATGCGCAGCTCAACAGCGGCG 360
OY 368 CAGAGAGCAATTCATTTGCTGATGACATGACAGAGATTTCACCAATTCGCCACCAAAAC 427
Db 361 CAGAGAGCAATTCATTTGCTGATGACATGACAGAGATTTCACCAATTCGCCACCAAAAC 420
OY 428 TGGCCGAGATCTTTGCTGCTGATCTCAGTCACAGTCCCTCAGTACAGTACAGTTGAGC 487
Db 421 TGGCCGAGATCTTTGCTGCTGATCTCAGTCACAGTCCCTCAGTACAGTACAGTTGAGC 480
OY 488 TGCATCTCAGACAGATAGCTGTGATGATGAGTTTCTCCCGAGAGAGAGCAAGCAAA 547
Db 481 TGCATCTCAGACAGATAGCTGTGATGATGAGTTTCTCCCGAGAGAGAGCAAGCAAA 540
OY 548 CTCGAAGGGCAGCAGCAATTTCTGTGTAAGAACATCAGACAGGCAATTTGGAGCGCC 607
Db 541 CTCGAAGGGCAGCAGCAATTTCTGTGTAAGAACATCAGACAGGCAATTTGGAGCGCC 600
OY 608 GAGATTTGAGTTGAGAGAGACAGATGCTGTGATTTGATTTGATTTGATTTGATTTG 667
Db 601 GAGATTTGAGTTGAGAGAGACAGATGCTGTGATTTGATTTGATTTGATTTGATTTG 660
OY 668 GGGGGAGAGCCCTTGGCTGTGCTTAAATAGTGGCTGTACACATCAGACAGCCAGAG 727
Db 661 GGGGGAGAGCCCTTGGCTGTGCTTAAATAGTGGCTGTACACATCAGACAGCCAGAG 720

```


QY 1440 AATATGGGCGACTCCAAACAGAAATGATGTACACGACCTCCGAGACTCCGAGCTGACG 1499
DB 181 AATATGGGCGACTCCAAACAGAAATGATGTACACGACCTCCGAGACTCCGAGCTGACG 240
QY 1500 TGGGAGCGAGTACGTTCTCTAGTGTGACATGTATCTGCGCAGATGGCAACGAGTGTG 1559
DB 241 TGGGAGCGAGTACGTTCTCTAGTGTGACATGTATCTGCGCAGATGGCAACGAGTGTG 300
QY 1560 CTCCTGGCAGAGGTCGCTCTACTCAATTTGAGCTGTCCACAGTTCACCTTTGTTCTG 1619
DB 301 CTCCTGGCAGAGGTCGCTCTACTCAATTTGAGCTGTCCACAGTTCACCTTTGTTCTG 360
QY 1620 TCCATCACACCCACAAACAGAGCTTTGGCACTGTATTAACCTCTAATGACCCGAGGGG 1679
DB 361 TCCATCACACCCACAAACAGAGCTTTGGCACTGTATTAACCTCTAATGACCCGAGGGG 420
QY 1680 CGATACAGAGAGATGTGTCTCTCTCTAAGAAATGATGAATAGTGCAGCTTG 1739
DB 421 CGATACAGAGAGATGTGTCTCTCTCTAAGAAATGATGAATAGTGCAGCTTG 480
QY 1740 CATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTG 1799
DB 481 CATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTG 540
QY 1800 GGACTCAACAAAAATGGGCCATTCAAACCTAATTTACAGATACTAATGACCATACTA 1859
DB 541 GGACTCAACAAAAATGGGCCATTCAAACCTAATTTACAGATACTAATGACCATACTA 600
QY 1860 CCAAGGACGAGTCCAGCTGACACACACACTTAAGAAATTTTAAAGATACTTTA 1919
DB 601 CCAAGGACGAGTCCAGCTGACACACACTTAAGAAATTTTAAAGATACTTTA 660
QY 1920 TTTTCTCTTACTCTCTCTCTCTGATTTTTCCTAATTTCAATTCCTTTTTCAT 1979
DB 661 TTTTCTCTTACTCTCTCTCTCTGATTTTTCCTAATTTCAATTCCTTTTTCAT 720
QY 1980 CTCATTAATCCAAG-TTCTGACAGACACAGAACTGTCTCATGCTC-TTAAATGAA 2037
DB 721 CTCATTAATCCAAGTTCTGACAGACACAGAACTGTCTCATGCTC-TTAAATGAA 780
QY 2038 ATAGAATGCA-GGGTCCCTCACTCTAGTCACTAAGAGATTTTAC-TCCCCAGGCC 2095
DB 781 ATAGAATGCAAGGGTCCCTCACTCTAGTCACTAAGAGATTTTACTTCTCCAGGCC 840
QY 2096 AGAAGGTGATTTCTCTTACATTTGCGG--ACTTAGCTTAA-TTAGGTACT 2152
DB 841 GAAAGGGGATTTCTCTTACATTTGCGGGAATTTAGCTTAAAGTACTACT 900
QY 2153 TATTACAGGAAAT-GCTAAGTACTCTCTGTGAA--CAATCTGCAATGTCTAAT 2208
DB 901 TATTACAGGAAATGTCTAAGTACTCTCTGTGAAACAAATCTGATGTCTAAT 960
QY 2209 CGCCTTAAAGAGCCCTTCTT 2231
DB 961 CGCCTTAAAGAGCCCTTCTT 983

Search completed: April 20, 2003, 17:23:39
Job time : 3309 secs

PL HATE DND;

XX
DR MPi; 1998-240074/21.
DR P-PsDB; AAM56097.
XX
PT New isolated S-adenosyl-L-homocysteine hydrolase enzyme - is used to
PT develop products which can be used in the treatment of e.g.
PT auto-immune disease, transplantations or cancers
XX
PS
XX Claim 7; Fig 1; 33pp; English.
XX
XX This is the nucleotide sequence of the DD4B9.3 enzyme which has
CC S-adenosyl-L-homocysteine hydrolase (AHcy)-type activity. Its
CC products can be used to identify substances which have a stimulatory
CC or inhibitory effect on the enzyme activity. Such substances can be
CC used to modulate dendritic cell (DC) function and for immunomodulation.
CC They can be used in the treatment of e.g. autoimmune diseases,
CC transplantsations or cancers. The products can also be used for
CC detection and diagnosis.
XX
XX Sequence 2563 BP; 646 A; 604 C; 677 G; 636 T; 0 other;

Query Match	100.0%	Score 417;	DB 19;	Length 2563;
Best Local Similarity	100.0%	Pred. NO.	2.2e-121;	
Matches 417;	Conservative 0;	Mismatches 0;	Gaps 0	

OY	1	GAGGAAACACCAACCAACCACTCCAAAGGGACACACATTTTCTGTGTAAACATCAAGC	60
Db	529	GAGAAACACCAACCAACCACTCCAAAGGGACACACATTTTCTGTGTAAACATCAAGC	588
OY	61	AGGCAGAAATTTGGACGCCGGGAGATTGAGATTGCAGAGCAAGACATCTCTCTGATTT	120
Db	589	AGGCAGAAATTTGGACGCCGGGAGATTGAGATTGCAGAGCAAGACATCTCTCTGATTT	648
OY	121	CACATCAGGAAACGTCCTCAGGGGGAGAAACCCCTTGCTGCTAAAAATGTGGGCTGTA	180
Db	649	CACATCAGGAAACGTCCTCAGGGGGAGAAACCCCTTGCTGCTAAAAATGTGGGCTGTA	708
OY	181	CACATCATACAGCCACAGCCAGCGGTGTGTAATGTGACACTGTGGCCCTGGGGGGCTAGT	240
Db	709	CACATCATACAGCCACAGCCAGCGGTGTGTAATGTGACACTGTGTGCCCTGGGGGGCTAGT	768
OY	241	GCCGCTGGTCTGCTTGTTAACATCTACTCAACTCAGAAATGAATGATGCACACTGGCTG	300
Db	769	GCCGCTGGTCTGCTTGTTAACATCTACTCAACTCAGAAATGAATGATGCACACTGGCTG	828
OY	301	AGGCTGAGATTGCAGTGTTCGCTTGGAAAGGCGAGTCAGAAAGATGACTTCTGTGTGTGA	360
Db	829	AGGCTGAGATTGCAGTGTTCGCTTGGAAAGGCGAGTCAGAAAGATGACTTCTGTGTGTGA	888
OY	361	TTGACCGCTGTGTAAACATGATGGTGGGAGCGACCAACATGATCCCGAGATGAGGGG	417
Db	889	TTGACCGCTGTGTAAACATGATGGTGGGAGCGACCAACATGATCCCGAGATGAGGGG	945

RESULT 2	
AAV73924	
ID	AAV73924 standard; DNA; 3616 BP.

AC	AAV73924;
XX	
DT	04-MAR-1999 (first entry)

DE Human SAHH DNA #1.

KM S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;
KM treatment; infection; cancer; autoimmune disease; detection; diagnosis;
KM gene mapping; antisense; therapy; antagonist; immunoassay; ss.

Homo sapiens

	Location/Qualifiers
EH	56..1558
FT	
FT.	/*tag= a
Key	
CDS	

FT		/product= "SAHH"
XX		
PM	.US5854023-A.	
XX		
PD	29-DEC-1998.	
XX		
PE	17-JUL-1997;	97US-0896005.
XX		
PR	17-JUL-1997;	97US-0896005.
XX		
PA	(INCY-) INCYTE PHARM INC.	
PI		
PI	Cortley NC, Hillman JL, Lal P, Shah P;	
XX		
DR	WPI; 1999-094906/08.	
DR	P-PSDS; AAW90061.	
XX		
PT	Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase	
PT	for production of recombinant enzyme, useful for diagnosis,	
PT	treatment and prevention of cancers, infections and autoimmune	
PT	diseases	
PS	Disclosure; Fig 1A-I; 40pp; English.	

This sequence encodes a human S-adenosyl-5-homocysteine hydrolase (SAHH). The SAHH protein can be used to generate specific antibodies and in drug screening to identify specific binding agents. Antagonists of the protein are used to treat or prevent a wide range of viral, bacterial, fungal, parasitic, protozoal or helminthic infections, many cancers (leukemia, lymphoma or solid tumours), and many autoimmune diseases (e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes mellitus, multiple sclerosis etc). All these conditions may be treated by expressing antisense sequences, triplex-forming agents or ribozymes directed against the nucleic acid. The nucleic acid and its fragments can be used as probes or primers for detecting and quantifying gene expression, for diagnosis or monitoring of disease, to identify genetic variations, mutations or polymorphisms, in gene mapping and as antisense therapeutics. Antibodies are used directly as antagonists, indirectly to deliver active agents to SAHH-expressing cells, to diagnose and monitor diseases in standard immunoassays, in competitive drug screens and to isolate the protein from natural sources.

..... SQ sequence 3616 BP; 1017 A; 782 C; 826 G; 991 T; 0 other;

Query Match	100.0%	Score 417;	DB 20;	Length 3616;
Best Local Similarity	100.0%	Pred. NC 2.6e-121;		
Matches 417;	Conservative 0;	Mismatches 0;	Gaps 0;	

Dy 1 GAGAAACGACGAACCACTCCAAAGGGCAGCAGCAATTTCTGTGTGAAGAACATCAAAGC 60
 |||||
 |||||
Db 240 GAGAAAGCGAACAACCACTCCAAGGGCAGCAGCAATTTCTGTGTGAAGAACATCAAAGC 29

Db 300 AGCGAATTGGAGCGCCGGGAGATTGAGATTGCGACGACACATGTGTCTGCTTGATTT 35

07 121 CACTAGGGAACGTCTCTAGGGGGGAGAACCCCTTGGCTGGTCTAAATTAAGTGGGCTGTA 18

Db 360 CACTAGGGAACGTCTCTAGGGGGGAGAACCCCTTGGCTGGTCTAAATTAAGTGGGCTGTA 41

QY 181 CACACATCACAGCCACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 24
|||||
Db 420 CACACATCACAGCCACAGCGGTGTTGATTGAGACACTCTGTGCCCTGGGGCTCAGT 47
|||||

241 GCCCGTGGTCTGCTTGTAACTACTCAACTCAGATGAATAGCTGCAGCACTGGCGTGG
 480 GCCCGTGGTCTGCTTGTAACTACTCAACTCAGATGAATAGCTGCAGCACTGGCGTGG
 533

301 AGCGTGAGTTGCACTGTTCCCTTGGAAAGGGCGAGTCAGAAGATGACTTCTGGTGGTGA 360

Db 540 AGGCTGGAGTTTCAGTGTTCGCTTGGAGGGCGAGTCAGAAATGACTTCTGGTGTGA 59

Db 600 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAACATGATCTGTGATGATGGG 656

RESULT 3

ABV23195
ID ABV23195 standard; cDNA: 3634 BP.

AC ABV23195;

PT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 23186.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US05171.

PR 17-FEB-2000; 2000US-183319P.

PR 16-MAR-2000; 2000US-189862P.

PR 25-MAY-2000; 2000US-207454P.

PR 09-JUN-2000; 2000US-211314P.

PR 18-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

PN WPI: 2001-662795/76.

PS Claim 1; Page 4168-4169; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate
cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a
patient;

CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence: 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;

Query Match 100.0%; Score 417; DB 23; Length 3634;

Best Local Similarity 100.0%; Pred. No. 2,6e-121;

Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GAGAGACGACGACCAACCACTCCAGAGGAGAGAGCAATTTCTGTGGAAGAACATCAAGC 60
599 GAGAGACGACGACCAACCACTCCAGAGGAGAGAGCAATTTCTGTGGAAGAACATCAAGC 658
61 AGGCAAGATTGGAGCGCGGGAGATTGAGATTCAGAGCAAGACATGCTGCTCTAATT 120
659 AGGCAAGATTGGAGCGCGGGAGATTGAGATTCAGAGCAAGACATGCTGCTCTAATT 718

QY 121 CACTCAGAGAAACGTGCTCAGAGGGAGAGAGCCCTTGCTGCTAAATAGTGGCTGTA 180
Db 719 CACTCAGAGAAACGTGCTCAGAGGGAGAGAGCCCTTGCTGCTAAATAGTGGCTGTA 778
QY 181 CACACATCAGAGCCGAGACAGCGGTGTGATTTAGACACATCTGTGCGCTGGGCTAGT 240
Db 779 CACACATCAGAGCCGAGACAGCGGTGTGATTTAGACACATCTGTGCGCTGGGCTAGT 838
QY 241 GCCGCTGCTGCTTGAACATCTACTCACTCAGATGAAGTGTGACACATGGCTG 300
Db 839 GCCGCTGCTGCTTGAACATCTACTCACTCAGATGAAGTGTGACACATGGCTG 898
QY 301 AGGCTGAGTTGAGAGTTCGCTTGAAGGCGAGTCAGAGATGACTTGTGTGTGTA 360
Db 899 AGGCTGAGTTGAGAGTTCGCTTGAAGGCGAGTCAGAGATGACTTGTGTGTGTA 958
QY 361 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAACATGATCTGTGATGGG 417
Db 959 TTGACCGCTGTGTGAACATGATGGTGGCAGGCCAACATGATCTGTGATGGG 1015

RESULT 4

ABV29032
ID ABV29032 standard; cDNA: 3634 BP.

AC ABV29032;

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 29023.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US05171.

PR 17-FEB-2000; 2000US-183319P.

PR 16-MAR-2000; 2000US-189862P.

PR 25-MAY-2000; 2000US-207454P.

PR 09-JUN-2000; 2000US-211314P.

PR 18-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

PN WPI: 2001-662795/76.

PS Claim 1; Page 6147-6148; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate
cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;

Query Match 100.0%; Score 417; DB 23; Length 3634;

Best Local Similarity 100.0%; Pred. No. 2,6e-121;

Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAGCAGCAACCACTCCAGGAGCAGCAGCAATTCCTGTGTAAGACATCAAGC 60
 DB 599 GAGAGAGCAGCAACCACTCCAGGAGCAGCAGCAATTCCTGTGTAAGACATCAAGC 658
 QY 61 AGGCAATTTTGGACCCCGGAGATTGAGATTGCAGAGCAATGCTGCTCTGATT 120
 DB 659 AGGCAATTTTGGACCCCGGAGATTGAGATTGCAGAGCAATGCTGCTCTGATT 718
 QY 121 CACTAGGAACGCTGCTCAGGGGAGAGCCCTGGCTGCTTAAATAGTGGCTGTA 180
 DB 719 CACTAGGAACGCTGCTCAGGGGAGAGCCCTGGCTGCTTAAATAGTGGCTGTA 778
 QY 181 CACACATCAGAGCCAGACAGCGGCTTGTGATGAGACACTGTGCTGGGGGCTCAGT 240
 DB 779 CACACATCAGAGCCAGACAGCGGCTTGTGATGAGACACTGTGCTGGGGGCTCAGT 838
 QY 241 GCGCGTGTCTGCTTGAACATCTACTCACTAGAAATGAGTGCAGCACTGGCTG 300
 DB 839 GCGCGTGTCTGCTTGAACATCTACTCACTAGAAATGAGTGCAGCACTGGCTG 898
 QY 301 AGGCTGAGTTGAGATGCTGCTTGAAGGGCAGTCAAGATGACTTGTGGTGTGA 360
 DB 899 AGGCTGAGTTGAGATGCTGCTTGAAGGGCAGTCAAGATGACTTGTGGTGTGA 958
 QY 361 TTGACCGCTGTGTGAACATGATGATGGTGGCAGGCCAATGATCTGGATGATGGG 417
 DB 959 TTGACCGCTGTGTGAACATGATGATGGTGGCAGGCCAATGATCTGGATGATGGG 1015

RESULT 5

AA576216/c

ID AA576216 standard; cDNA; 5030 BP.

XX AA576216;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #12020.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

XX 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Dmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR P-PSDB; ABG12029.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

PS Claim 1; SEQ ID No 12020; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations/
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AA564197-AA594564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 5030 BP; 1432 A; 1137 C; 1178 G; 1283 T; 0 other;

Query Match 65.9%; Score 274.6; DB 23; Length 5030;

Best Local Similarity 78.7%; Pred. No. 3.3e-76;

Matches 328; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1 GAGAGAGCAGCAACCACTCCAGGAGCAGCAGCAATTCCTGTGTAAGACATCAAGC 60
 DB 4489 GAGAGAGCAGCAACCACTCCAGGAGCAGCAGCAATTCCTGTGTAAGACATCAAGC 4430
 QY 61 AGGCAATTTTGGACCCCGGAGATTGAGATTGCAGAGCAATGCTGCTGATTT 120
 DB 4429 AGGCAATTTTGGACCCCGGAGATTGAGATTGCAGAGCAATGCTGCTGATTT 4370
 QY 121 CACTCAGGAACGCTGCTCAGGGGAGAGCCCTGGCTGCTTAAATAGTGGCTGTA 180
 DB 4369 CTTTGAAGAGAGAGCTCAAGAGAAAGACCTTGGCTGGAGCCAAATCTGGGTTGCA 4310
 QY 181 CACACATCAGAGCCAGACAGCGGCTTGTGATGAGACACTGTGCTGGGGGCTCAGT 240
 DB 4309 CACACATCAGAGCCAGACAGCGGCTTGTGATGAGAACTCTGGGCTGGGGCCAGT 4250
 QY 241 GCGCGTGTCTGCTTGAACATCTACTCACTAGAAATGAGTGCAGCACTGGCTG 300
 DB 4249 GCGCGTGTCTGCTTGAACATCTACTCACTAGAAATGAGTGCAGCACTGGCTG 4150
 QY 301 AGGCTGAGTTGAGATGCTGCTTGAAGGGCAGTCAAGATGACTTGTGGTGTGA 360
 DB 4189 AAAGTGATTTTCCCTGTTTGGCTGGAGAGGAGTCAAGATGACTTGTGGTGTGA 4130
 QY 361 TTGACCGCTGTGTGAACATGATGATGGTGGCAGGCCAATGATCTGGATGATGGG 417
 DB 4129 TTGACCGCTGTGTGAACATGATGATGGTGGCAGGCCAATGATCTGGATGATGGG 4073

RESULT 6

ABV44141/c

ID ABV44141 standard; cDNA; 504 BP.

XX ABV44141;

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 44132.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KM pharmacogenomic marker; gene; ss.
XX Homo sapiens.
OS
XX MO200160860-A2.
PN
XX 23-AUG-2001.
PD
XX
PF 20-FEB-2001; 2001MO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-235281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
DR
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer. useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 8770; 11750pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 504 BP; 117 A; 150 C; 116 G; 121 T; 0 other;
SQ
Query Match 57.6%; Score 240; DB 23; Length 504;
Best Local Similarity 100.0%; Pred. No. 1e-65;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 178 GTACACATCAGAGCCAGAGCGGTTGATGAGACACTCTGCGCCCTGGGGGCTC 237
DB 502 GTACACATCAGAGCCAGAGCGGTTGATGAGACACTCTGCGCCCTGGGGGCTC 443
XX
QY 238 AGTCCCGTGTCTGCTTGTAACTCTACTCACTCAAGATGAAGTAGCTGCAGACCTGG 297
DB 442 AGTCCCGTGTCTGCTTGTAACTCTACTCACTCAAGATGAAGTAGCTGCAGACCTGG 383
XX
QY 298 CTGAGCGTGTGAGTGTGCTTCTGTTGGAAGCGGAGTCAAGAAATGACTCTGCTGTGT 357
DB 382 CTGAGCGTGTGAGTGTGCTTCTGTTGGAAGCGGAGTCAAGAAATGACTCTGCTGTGT 323
XX
QY 358 GTATTGACCGCTGTGTGAACATGATGGTGGCAGGCCAATCATGATCCGTGATGGGG 417
DB 322 GTATTGACCGCTGTGTGAACATGATGGTGGCAGGCCAATCATGATCCGTGATGGGG 263
XX
RESULT 7
AAZ80766/c
ID AAZ80766 standard; cDNA; 636 BP.
XX
AC AAZ80766;

XX
XX 07-APR-2000 (first entry)
DT
XX Human colon cancer cell line SW480 cDNA clone SEQ ID NO:850.
DE
XX
XX Human; gene expression product; diagnosis; tumour; colon cancer;
KM colorectal adenocarcinoma; cell line SW480; cell proliferation;
KM cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
KM hyperplasia; ds.
XX
XX Homo sapiens.
OS
XX
XX MO964576-A2.
PN
XX 16-DEC-1999.
PD
XX
XX 09-JUN-1999; 99MO-IB01062.
PF
XX
XX 10-JUN-1998; 98US-0088801.
PR
XX
XX (FARB) BAYER CORP.
PA
XX
XX Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;
PI Carroll E, Catino TJ, Dertl A, Ford DM, Lewis ME, Monahan JE;
PI Schlegel R.
XX
XX WPI; 2000-087220/07.
DR
XX
XX Novel nucleic acids, used to develop products for the diagnosis and
PT treatment of disorders involving unwanted cell proliferation,
PT particularly cancers, especially colon cancer -
XX
XX Claim 15; Page 469; 469pp; English.
PS
XX
XX AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from
CC the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
CC cDNA clones can be used to generate antisense oligonucleotides which
CC can be used for antisense therapy. Methods and products from the present
CC invention can be used for identifying and/or classifying cancerous cells
CC present in a human tumour, particularly in solid tumours, e.g.
CC carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
CC can be used for developing agents for the diagnosis and treatment of
CC disorders involving unwanted cell proliferation, such as neoplasia,
CC dysplasia or hyperplasia.
XX
XX Sequence 636 BP; 156 A; 166 C; 137 G; 147 T; 30 other;
SQ
Query Match 54.3%; Score 226.4; DB 21; Length 636;
Best Local Similarity 99.2%; Pred. No. 2.3e-61;
Matches 238; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
XX
QY 179 TACACATCAGAGCCAGAGCGGTTGATGAGACACTCTGCGCCCTGGGGGCTC 237
DB 420 TACACATCAGAGCCAGAGCGGTTGATGAGACACTCTGCGCCCTGGGGGCTC 361
XX
QY 238 AGTCCCGTGTCTGCTTGTAACTCTACTCACTCAAGATGAAGTAGCTGCAGACCTGG 297
DB 360 AGTCCCGTGTCTGCTTGTAACTCTACTCACTCAAGATGAAGTAGCTGCAGACCTGG 301
XX
QY 298 CTGAGCGTGTGAGTGTGCTTCTGTTGGAAGCGGAGTCAAGAAATGACTCTGCTGTGT 357
DB 300 CTGAGCGTGTGAGTGTGCTTCTGTTGGAAGCGGAGTCAAGAAATGACTCTGCTGTGT 241
XX
QY 358 GTATTGACCGCTGTGTGAACATGATGGTGGCAGGCCAATCATGATCCGTGATGGGG 417
DB 240 GTATTGACCGCTGTGTGAACATGATGGTGGCAGGCCAATCATGATCCGTGATGGGG 181
XX
RESULT 8
ABV35309
ID ABV35309 standard; cDNA; 312 BP.
XX
AC ABV35309;

XX 16-SEP-2002 (first entry)
XX Human prostate expression marker CDNA 35300.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US05171.
XX 17-FEB-2000; 2000US-183319P.
XX 16-MAR-2000; 2000US-189862P.
XX 25-MAY-2000; 2000US-207454P.
XX 09-JUN-2000; 2000US-211314P.
XX 18-JUL-2000; 2000US-219007P.
XX 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 7360-7361; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 312 BP; 69 A; 71 C; 102 G; 70 T; 0 other;
XX
XX Query Match 54.1%; Score 225.8; DB 23; Length 312;
XX Best Local Similarity 98.8%; Pred. No. 2.6e-61;
XX Matches 238; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
XX
XX 178 GTACACATCAGACAGCCAGAGAGCGGTGTGATGAGACACTGTCGCGGGGGCTC 237
XX |||||
XX DB 53 GTACACATCAGACAGCCAGAGAGCGGTGTGATGAGACACTGTCGCGGGGGCTC 112
XX
XX 238 AGTCCGCTGCTGCTGTAACTACTCACTCAGATGAAGTGCAGACTGG 297
XX |||||
XX DB 113 AGTGTGCTGCTGCTGTAACTACTCACTCAGATGAAGTGCAGACTGG 172
XX
XX 298 CTGAGGCTGAGTGTGAGTGTGCTTGAAGGGCAGTCAAGATG-ACCTTCTGGTGG 356
XX |||||
XX DB 173 CTGAGGCTGAGTGTGAGTGTGCTTGAAGGGCAGTCAAGATGAAGTCTTCTGGTGG 232
XX
XX 357 TGTATTGACCGGTGTGATGAATGATGGTGGAGCCCAACATGATCCGTGATGATGG 416
XX |||||
XX DB 233 GGTATTGACCGGTGTGATGAATGATGGTGGAGCCCAACATGATCCGTGATGATGG 292

QY 417 G 417
DB 293 G 293
RESULT 9
ABV14215
ID ABV14215 standard; CDNA; 435 BP.
XX
XX AC ABV14215;
XX
XX 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker CDNA 14206.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US05171.
XX 17-FEB-2000; 2000US-183319P.
XX 16-MAR-2000; 2000US-189862P.
XX 25-MAY-2000; 2000US-207454P.
XX 09-JUN-2000; 2000US-211314P.
XX 18-JUL-2000; 2000US-219007P.
XX 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 2369-2370; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 435 BP; 107 A; 86 C; 133 G; 109 T; 0 other;
XX
XX Query Match 54.1%; Score 225.8; DB 23; Length 435;
XX Best Local Similarity 98.8%; Pred. No. 3e-61;
XX Matches 238; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
XX
XX 178 GTACACATCAGACAGCCAGAGAGCGGTGTGATGAGACACTGTCGCGGGGGCTC 237
XX |||||
XX DB 3 GTACACATCAGACAGCCAGAGAGCGGTGTGATGAGACACTGTCGCGGGGGCTC 62
XX
XX 238 AGTCCGCTGCTGCTGTAACTACTCACTCAGATGAAGTGCAGACTGG 297

Db 63 AGTGGCTGCTGCTGCTGTAACATCTACTCACTCAAGTAAGTAAGTACGACGACTGG 122

QY 298 CTGAGCGTGGAGTGTGCAATGTTCCCTTGGAAGGGGAGTCAAGAAATG-ACCTTCGGTGG 356

Db 123 CTGAGCGTGGAGTGTGCAATGTTCCCTTGGAAGGGGAGTCAAGAAATGAACTTCGGTGG 182

QY 357 TGTATTGACCGCTGTGTGATGATGGTGGCGAGGCAACATGATCTGATGATGGG 416

Db 183 GGTATTGACCGCTGTGTGATGATGGTGGCGAGGCAACATGATCTGATGATGGG 242

QY 417 G 417

Db 243 G 243

RESULT 10

ABL03693

Id ABL03693 standard; cDNA; 1566 BP.

XX

AC ABL03693;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 5561.

XX

KM Drosophila; developmental biology; cell signalling; insecticide.

KW pharmaceutical; gene; ss.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PMD, Myers EW;

XX

DR WPI; 2001-656860/75.

XX

DR P-PSDB; ABB59590.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX

PS Claim 1; SEQ ID NO 5561; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

CC sequences (AB101840-AB16175) and the encoded proteins

CC (ABB5737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 1566 BP; 360 A; 438 C; 481 G; 287 T; 0 other;

QY Query Match 50.18; Score 208.8; DB 23; Length 1566;

Best Local Similarity 70.28; Pred. No. 1.3e-55;

Matches 295; Conservative 0; Mismatches 122; Indels 3; Gaps 1;

QY 1 GAGAGAGCAACCAACCACTCCAGGCGAGCAGCAATTTCTGTGAGAAACAT---CA 57

Db 245 GAGAGAGCTCCAAAGAACTCCAAAGCGAGCTCGATTCTGTGTAAGAAACATGCGG 304

QY 58 AGCAGCAGAAATTTGGAGCGCGGAGATTGAGATTGACAGAGACATGTCGCTGTA 117

Db 305 CCCAGATCATTTTGGGCCAGGAGATCGAGATTTCGGAGCAGAGATGCGGCGATCA 364

QY 118 TTTCATCTAGAGAAACGCTGCTCAGGGGAGAGAGCCCTTGGCTGCTAAATAGTGGCT 177

Db 365 TAGCGCTGAAGAGCGGGGCGCGAGACAGCTTGTGAAGATGTCGAAGATGCTGGGT 424

QY 178 GTACACATCACAGCCAGACAGCGGTGTGATTGACACTCTGTGCCCTGGGGCTC 237

Db 425 GCACACATCAACGCCAGACGCGCTGCTCATCAGAGCGGTGTGAACTGGGGCGCA 484

QY 238 AGTGGCGTGGTCTGCTGTAACATCTACTCACTCAAGTAAGTAAGTACGACGACTGG 297

Db 485 GCGTGGCGTGGCGCGCTGCAACATCTACTCCAGCAGAGAGAGTGGCTGCTTGG 544

QY 298 CTGAGCGTGGAGTGTGCAATGTTCCCTTGGAAGGGGAGTCAAGAAATGACTTGGTGGT 357

Db 545 CCGAATCCGGAATCCCGATCTTGCCTGCGCGCGCAGACGAGAGAGACTTGTGGTGT 604

QY 358 GTATTGACCGCTGTGTGAACATGATGGTGGCGAGGCCAACATGATCTGATGATGGG 417

Db 605 GCATAGATGCTGCTGTAACGCGGAGAACTGCGAGCCAAACATGATCTGAGATGGG 664

RESULT 11

ABL29787

Id ABL29787 standard; DNA; 1776 BP.

XX

AC ABL29787;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 40834.

XX

KM Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PMD, Myers EW;

XX

DR WPI; 2001-656860/75.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX

PS Claim 1; SEQ ID NO 40834; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

CC sequences (AB101840-AB16175) and the encoded proteins

CC (ABB5737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX AC ABV05046;
XX 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 5037.
XX
KW Human: prostate cancer: cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-18319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
XX WPI: 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer; useful
PT for detecting presence of prostate cancer; stage of prostate cancer -
XX
PS Claim 1; Page 857; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 328 BP; 77 A; 67 C; 102 G; 77 T; 5 other;
XX
Query Match 37.6%; Score 156.8; DB 23; Length 328;
Best Local Similarity 89.6%; Pred. No. 1.8e-39;
Matches 225; Conservative 0; Mismatches 15; Indels 11; Gaps 5;
XX
QY 178 GTACACATCATCAGCCAGACA-GGGGTGTTGATGAGACACTGTGCGCTGGGGCT 236
DB 19 GTACACATCATCAGCCAGACA-GGGGTGTTGATGAGACACTGTGCGCTGGGGCT 78
QY 237 CAGTGCCTGCTGCTCTGTACATCTCACTCACTCAGATGAGTAGTGCAGCACTG 296
DB 79 CAGTGCCTGCTGCTCTGTACATCTCACTCACTCAGATGAGTAGTGCAGCACTG 138
QY 297 GCTGAGCTGGAAGTTCAGTTCGCTT-GGAAGGCGAGTCAAGATGACTTC---- 350
DB 139 GCTGAGCTGGAAGTTCAGTTCGCTT-GGAAGGCGAGTCAAGATGACTTCGCTG 198
QY 351 TGGTGTGATTTACCGCTGTGTG--ACATGATGGGTGG--CAGGCCAATGATCCT 406

DB 199 TGGTGTGATTTACCGCTGTGTGGAACCTGTGTGGAACATTTGATGGTNGGCAAGCCACATGATCCT 258
QY 407 GGATGATGGGG 417
DB 259 GGATGATGGGG 269

Search completed: April 21, 2003, 00:05:25
Job time : 97.6286 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 22:06:49 ; Search time 1076.44 Seconds

(without alignments)
11274.088 Million cell updates/sec

Title: US-09-782-051-1_COPY_529_945
Page: 417

Sequence: 1 gagaggaagcagcaaccac.....catgatcctgatgatggg 417

Scoring table: IDENTITY_NTUC Count 1 0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database

```

1:  GenEmb1:*
2:      gb_pa:*
3:      gb_htg:*
4:      gb_in:*
5:      gb_cm:*
6:      gb_ov:*
7:      gb_pat:*
8:      gb_ph:*
9:      gb_pl:*
10:     gb_pr:*
11:     gb_ro:*
12:     gb_sts:*
13:     gb_sy:*
14:     gb_un:*
15:     gb_vl:*
16:     em_da:*
17:     em_fun:*
18:     em_hum:*
19:     em_in:*
20:     em_mu:*
21:     em_or:*
22:     em_ov:*
23:     em_pat:*
24:     em_ph:*
25:     em_pl:*
26:     em_ro:*
27:     em_sts:*
28:     em_un:*
29:     em_vl:*
30:     em_htg_hum:*
31:     em_htg_inv:*
32:     em_htg_other:*
33:     em_htg_mus:*
34:     em_htg_pln:*
35:     em_htg_rtd:*
36:     em_htg_mam:*
37:     em_htg_vrt:*
38:     em_sy:*
39:     em_htgo_hum:*
40:     em_htgo_mus:*
41:     em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
	1	417	100.0	2258	9	HS082761	U82761	Homo sapien
	2	417	100.0	2526	9	BC007576	BC007576	Homo sapi
	3	417	100.0	2552	9	BC016942	BC016942	Homo sapi
	4	417	100.0	2563	6	AX029176	AX029176	Sequence
	5	417	100.0	2583	9	BC010681	BC010681	Homo sapi
	6	417	100.0	2677	9	AF315687	AF315687	Homo sapi
	7	415.4	99.6	2510	9	HS800298	HS800298	Homo sapi
	8	385	62.3	1772	10	BC018218	BC018218	Homo sapi
	9	276.2	66.2	1987	9	AK025372	AK025372	Mus muscu
	10	274.6	65.9	2030	9	BC008349	BC008349	Homo sapi
	11	274.6	65.9	5025	9	AB020635	AB020635	Homo sapi
	12	274.6	65.9	5052	9	BC024325	BC024325	Homo sapi
	13	191	45.8	1792	3	AY113501	AY113501	Drosophila
	14	191	45.8	52423	2	AC020395	AC020395	Drosophila
	15	191	45.8	80423	3	DROABDB	DROABDB	Drosophila
	16	191	45.8	175335	3	AC091636	AC091636	Drosophila
	17	191	45.8	223098	3	AE003715	AE003715	Drosophila
	18	191	45.8	338234	3	DM031961	DM031961	Drosophila
	19	189.4	45.4	5050	3	DMEX200	DMEX200	Drosophila
	20	144.2	34.6	2057	10	MGSS5MH	MGSS5MH	Mus musculus
	21	144.2	34.6	2067	10	BC015304	BC015304	Mus musculus
	22	144.2	34.6	190612	2	AC129093	AC129093	Mus muscu
	23	139.4	33.4	2029	10	RATAMHA	RATAMHA	Rat s-denno
	24	138.8	33.3	2211	6	ARI44978	ARI44978	Sequence
	25	138.8	33.3	2211	6	AX050466	AX050466	Sequence
	26	138.8	33.3	2211	6	AX069338	AX069338	Sequence
	27	138.8	33.3	2211	9	HDMRHCY	HDMRHCY	Human S-ade
	28	138.8	33.3	2355	9	AK097610	AK097610	Homo sapi
	29	137.8	33.0	1302	6	AX458351	AX458351	Sequence
	30	137.2	32.9	2084	6	HDMRHCY2	HDMRHCY2	Human S-ade
	31	137.2	32.9	2175	9	BC010018	BC010018	Homo sapi
	32	137.2	32.9	2183	9	BC011606	BC011606	Homo sapi
	33	137	32.9	2088	5	XLA7835	XLA7835	Xenopus L
	34	136.2	32.7	1631	3	LETSADSH	LETSADSH	Xenopus L
	35	133.8	32.1	2129	5	XEL5AH	XEL5AH	L. donovani
	36	129.8	31.1	110000	2	LMFICR36_-08	LMFICR36_-08	Xenopus lae
	37	129.8	31.0	40320	2	AC015387	AC015387	Continuation (9 of
	38	129.2	31.0	173372	3	AC010110	AC010110	Continuation (9 of
	39	129	30.9	1564	6	DMAHCXGEN	DMAHCXGEN	Drosophila
	40	128.2	30.7	1000	6	AX081541	AX081541	Drosophila
	41	128.2	30.7	1000	6	AX374722	AX374722	Drosophila
	42	128.2	30.7	58909	6	AX067455	AX067455	Sequence
	43	127.4	30.6	1603	3	AY102668	AY102668	Sequence
	44	127.4	30.5	10910	1	AE004480	AE004480	Pseudomon
	45	124	29.7	12968	1	AE012843	AE012843	Chlorobiu

ALIGNMENTS

RESULT 1	
HSU82761	
LOCUS	2258 bp mRNA linear PRI 17-DEC-2001
DEFINITION	Homo sapiens S-adenosyl homocysteine hydrolase homolog (XPYkxona)
ACCESSION	mRNA, complete cds.
VERSION	U82761
KEYWORDS	U82761.1 GI:2852124
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	Cleaver,J.E., Afzal,V., Feeney,L., McDowell,M., Sadinski,W.,
AUTHORS	Volpe,J.P.G., Busch,D.B., Coleman,D.M., Ziffer,D.W., Yu,Y.,

FEATURES	REFERENCE	JOURNAL	AUTHORS	TITLE
source	1. 2258			Nagasawa, H. and Little, J. B.
				Increased ultraviolet sensitivity and chromosomal instability related to p53 function in the xeroderma pigmentosum variant
				Cancer Res. 59 (5), 1102-1108 (1999)
				99168517
				10070969
				2 (bases 1 to 2258)
				Volpe, J. P. G., McDowell, M., Jostes, R. F., Afzal, V., Sadinski, W.,
				Trask, B. J., Legerski, R. and Cleaver, J. E.
				Complementation of chromosomal instability in the xeroderma pigmentosum variant by a gene on human chromosome 1 with homology to S-adenosyl homocysteine hydrolase
				unpublished
				3 (bases 1 to 2258)
				Volpe, J. P. G., McDowell, M. and Cleaver, J. E.
				Direct Submission
				Submitted (19-DEC-1996) Dermatology, UCSF, 3rd and Parnassus, Box
				0750, San Francisco, CA 94143, USA
				Location/Qualifiers
				1. 2258
				/organism="Homo sapiens"
				/db_xref="taxon:9606"
				/chromosome="1"
				/cell_type="fibroblasts"
				/tissue_type="skin"
				1. 2258
				/gene="XPVKona"
				48. 1550
				/gene="XPVKona"
				/note="xeroderma pigmentosum variant"
				/codon_start=1
				/product="S-adenosyl homocysteine hydrolase homolog"
				/protein_id="AAC01960.1"
				/db_xref="GI:2852125"
				/translation="MATVTEKAPKKOIFADDMOEFTEPKRTGRSLRSISOSSTDS
				YSSASLSDSSDEVSPEKREKQNTSKSNMCYCNITKAOERGREITLARDMSALIS
				LRRKAOEKLPLAKGIYTCITHTTQAVLITLCAIQAQCKMSACNITSTONEVAAL
				AEAGVAFVAMKGESEDFMWCIDRCVMDGDQAMLLDDGDLTHWYKRYKPNFKRI
				RGVIESVFAVHRLTYQLSKPLCVPAMNVDVTKQKPNLYCCRESILDLKRTID
				VMGCGAQQVVCVGEVGEVKGCCAKALGAIVYITLIDPICALQMDGFRVYKLNEDY
				RDNDVVTTCGNKNVYTRHLDLRKMSNICYGNMHSMTEDVTSIRPELMEWRVRSO
				VDAHLPMDGRVVLVLAEGRLNLSCSVPIPFVLSITQTQALALIELYNAPEGRKOD
				VYLLPKMDERIVASLHLPSPDHAHTELTDOAKTLGINKGPFNPNTYR"
BASE COUNT	603 a	513 c	542 g	600 t
ORIGIN				
Query Match	100.0%;	Score 417;	DB 9;	Length 2258;
Best Local Similarity	100.0%;	Pred. No. 3.7e-117;		
Matches 417;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Y	1	GAGAGAACGACCAACCAACTCCAAAGGACAGCAGCAATTTCTGTGTGAAGACATCAAGC	60	
Db	232	GAGAGAACGACCAACCAACCAACTCCAAAGGACAGCAGCAATTTCTGTGTGAAGACATCAAGC	291	
Y	61	AGGAGAAATTTGGAGCGCCGGGAGATTGAGATTGCAAGACAAACATGCTGCTGTGATTT	120	
Db	292	AGGAGAAATTTGGAGCGCCGGGAGATTGAGATTGCAAGACAAACATGCTGCTGTGATTT	351	
Y	121	CACACGAGAAACGCGACAGGGGGAGAGAAAGCCCTGGCTGGGTGCTAAATATATGGGCTGTA	180	
Db	352	CACACGAGAAACGCGCTCAGGGGGAGAGAAAGCCCTGGCTGGGTGCTAAATATATGGGCTGTA	411	
Y	181	CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGGCCCTGGGGGCTCGT	240	
Db	412	CACACATCACAGCCAGACAGCGGTGTGATTGAGACACTGTGTGGCCCTGGGGGCTCGT	471	
Y	241	GCCGCTGCTGCTGTGTAACATCTACTCAACTCAGAAATGAAGTGAAGTGCACACTGGCTG	300	
Db	472	GCCGCTGCTGCTGTGTAACATCTACTCAACTCAGAAATGAAGTGAAGTGCACACTGGCTG	531	
Y	301	AGGCTGAGATGACAGTGTGCTGCTTGGAAAGGCGAGTCAAGAAATGACTTCTGTGGTGTGA	360	
Db	532	AGGCTGAGATGACAGTGTGCTTGGAAAGGCGAGTCAAGAAATGACTTCTGTGGTGTGA	591	

	Oy	361	TTCACCGCTGTGTTGAATCATGGTGGTGAGGCCAACATGATTCCMGATGATGGGG	417
Dd	592	TTGACCCTGTGTGTAACAATGCATGGATGGGTGGCAGCCCACACTGATCTCTGGATGATGGGG	648	

RESULT 2
LOCUS BC007576
DEFINITION Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone MGc:1558 IMAGE:3139729, mRNA, complete cds.
ACCESSION VERSION BC007576
KEYWORDS GI:14043176
SOURCE Homo saplens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo;
1 (bases 1 to 2526)
Strausberg,R.
Direct Submission
Submitted (10-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: gcgabs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN).
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chin, Chris Ejelli, Erin Garland, Ran Guin,
Letitia Hsiao, Martin Krzywinski, Reta Kutsche, Olivier Lee, Soo Sen Lee, Victor Ling, Carrie Matthews, Candace McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Seedei, Jacqueline Scheel, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLn at: http://image.lnl.nl.gov Series: IRAL Plate: 22 Row: h Column: 21 This clone was selected for full length sequencing because it passed the following selection criteria: hexamer frequency ORF analysis, similarity but not identity to protein. location/qualifiers
1..2526
 /organism="Homo saplens"
 /db_xref="locusid:10766"
 /db_xref="taxon:9606"
 /clone="MGC:1558 IMAGE:3139729"
 /tissue_type="Placenta, choriocarcinoma"
 /clone_id="NH_MGC_21"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
308..1810
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH07576.1"
/db_xref="GI:14043177"
/translation="MATVTKAPKKQIQFDADMOEFTRFPKRTGRSRSSRSISOSTDSYSSASRITSDSEDEVSREKROQTNSKSNNFCYNKIQAEGREIREIAODMSALILRRRAOGEFLAKAIIVGCHTIQAQVALLETICALCAOCRWSCNTYSTONEVAALLAEVGAFPAWKGSEDPEFWNICIDRCVMNDGMQAMMLDDGDGLTHMYKYPNVFKEIRGVESVTVEHLRYLOYSKAGKLCPVPMMNVSTYOKPNIKVCRESLTDLGKRITDVMGGGOVVYGCVGEIGKCAGCALIKPALAVNYTEIDPCIALOCAOGDFPVLYTVIROYDYVITCKGNKNYYTRIHLDRKSNLCYCNGHSHTEDYTSLRPETLRWRVNSQVDHYTFPGDKRVVLABGRLLTNCSVPPLFTTLTDALAILIEYNPERPRVNOVYLIPKMDYEVASLIHPSPDALTELTDQAVTYLGINKNKPKPYPTFR"
```

FEATURES SOURCE CDS

AUTHORS Hart, D.N.
Enzyme having s-adenosyl-L-homocysteine hydrolase (ahcy) type
activity
JOURNAL Patent: WO 9814562-A 1 09-APR-1998;
HART, DEREK NIGEL JOHN (NZ)
FEATURES
source
 1. 2563
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /protein_id="CAC09285.1"
 /note="Open reading frame extends without a stop codon for the full 5' nucleotide sequence. The initiation codon has yet to be identified."
 /codon_start=1
 /protein_id="CAC09285.1"
 /db_xref="GI:10190061"
 /translation="RGVGRASCCFSCFGRHRCPCALGCTRTDKAMATAPQRTQDLAGRVRGVRSEGAARAGRAPRERGGGSGAAGAGAGAMPAMPAPGVGEELKAKIEDAKYSFMATVTRAPKQIQFADDMQETKPTKGRSLRSISOSTSYSASAATDSDESPREKQTNKSGSNFCVKNIKQAEGRREIARQDMALISIRKRAGERIAGAKITGCTHITQATVLETCALACACRSCSNCTSTONETAAALAEAGAVFAFKGSEDEDFWNCIDRCVNMDSQAMNLLDDGDLTHWYKRYPNFKIRGIVESVTGVRHLYOLSKAGKLCVPMANVNSVTOKPFDNLXCCRESILDLKRTDVMFGKROVVYCGEYVGGCCALKALGAIYIETIDPICALQACMDGFRVAKNEYIQRQVDVITCTGNKNVTRREHLDNRKNSCIQNMGSNTEIDVTSLETPELTWERVSQVDHYVDPGRVYLAEGRLNLSCTVPTFVLSITATQALALIELYNAPESRYKQDYLPRKMDYVASLHLPSPFDALHLEFLTDQAKYLGNKNGPFRPNYR"

BASE COUNT 646 a 604 c 677 g 636 t

ORIGIN

Query Match 100.0%; Score 417; DB 6; Length 2563;
 Best Local Similarity 100.0%; Pred. No. 3.7e-117;
 Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGAGAGCAGCAAAACCACTCCAAAGGCGAGCAGCAATTCCTGTGTAAGAACATCAAGC 60
 |||||||
 DB 529 GAGAGAGCAGCAAAACCACTCCAAAGGCGAGCAGCAATTCCTGTGTAAGAACATCAAGC 588
 OY 61 AGCGAAGATTGGAGCGCGGAGATTGAGATTCGACAGCAACATGCTGCTGATTT 120
 |||||||
 DB 589 AGCGAAGATTGGAGCGCGGAGATTGAGATTCGACAGCAACATGCTGCTGATTT 648
 OY 121 CACTAGCAAGAGCTGCTCAGGGGAGAAAGCCCTGGCTGCTTAATAATAGTGGCTGTA 180
 |||||||
 DB 649 CACTAGCAAGAGCTGCTCAGGGGAGAAAGCCCTGGCTGCTTAATAATAGTGGCTGTA 708
 OY 181 CACACATCAACAGCCAGACAGAGCGGTGATTGAGACACTCTGTGCCCTGGGGGCTCACT 240
 |||||||
 DB 709 CACACATCAACAGCCAGACAGCGGTGATTGAGACACTCTGTGCCCTGGGGGCTCACT 768
 OY 241 GCGGCTGCTGCTGTTAATCACTCACTCACTCACTCACTCACTCACTCACTCACTG 300
 |||||||
 DB 769 GCGGCTGCTGCTGTTAATCACTCACTCACTCACTCACTCACTCACTCACTG 828
 OY 301 AGCGTGGAGTTGCACTGTTCCCTTGGAGAGGCGGAGTCAAAAGATGACTTCTGCTGTA 360
 |||||||
 DB 829 AGCGTGGAGTTGCACTGTTCCCTTGGAGAGGCGGAGTCAAAAGATGACTTCTGCTGTA 888
 OY 361 TTGACCGCTGTGTAACATGATGATGGTGGCAGGCAACATGATGATGATGATGATG 417
 |||||||
 DB 889 TTGACCGCTGTGTAACATGATGATGGTGGCAGGCAACATGATGATGATGATGATG 945

RESULT 5
 BC010681 2563 bp mRNA linear PRI 12-JUL-2001
LOCUS BC010681
DEFINITION Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
IMAGE:8936 IMAGE:3853747, mRNA, complete cds.
ACCESSION BC010681
VERSION BC010681.1 GI:14715037
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE
 1 (bases 1 to 2563)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA

REMARK
COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalobonbcm.tmc.edu
 Villalobon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRK Plate: 13 Row: b Column: 21
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA 91:5729723.
 Location/Qualifiers
 1. 2583
 /organism="Homo sapiens"
 /db_xref="locustid:10768"
 /db_xref="taxon:9606"
 /clone="MGC:8936 IMAGE:3853747"
 /tissue.type="Colon, adenocarcinoma"
 /clone.lib="NIH_MGC_65"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPOrt6"
 367. 1869
 /codon_start=1
 /product="S-adenosylhomocysteine hydrolase-like 1"
 /protein_id="AAH10681.1"
 /db_xref="GI:14715038"
 /translation="MATVTRAPKQIQFADDMQETKPTKGRSLRSISOSTSYSASAATDSDESPREKQTNKSGSNFCVKNIKQAEGRREIARQDMALISIRKRAGERIAGAKITGCTHITQATVLETCALACACRSCSNCTSTONETAAALAEAGAVFAFKGSEDEDFWNCIDRCVNMDSQAMNLLDDGDLTHWYKRYPNFKIRGIVESVTGVRHLYOLSKAGKLCVPMANVNSVTOKPFDNLXCCRESILDLKRTDVMFGKROVVYCGEYVGGCCALKALGAIYIETIDPICALQACMDGFRVAKNEYIQRQVDVITCTGNKNVTRREHLDNRKNSCIQNMGSNTEIDVTSLETPELTWERVSQVDHYVDPGRVYLAEGRLNLSCTVPTFVLSITATQALALIELYNAPESRYKQDYLPRKMDYVASLHLPSPFDALHLEFLTDQAKYLGNKNGPFRPNYR"

BASE COUNT 648 a 606 c 688 g 641 t

ORIGIN

Query Match 100.0%; Score 417; DB 9; Length 2583;
 Best Local Similarity 100.0%; Pred. No. 3.7e-117;
 Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGAGAGCAGCAAAACCACTCCAAAGGCGAGCAGCAATTCCTGTGTAAGAACATCAAGC 60
 |||||||
 DB 551 GAGAGAGCAGCAAAACCACTCCAAAGGCGAGCAGCAATTCCTGTGTAAGAACATCAAGC 610
 OY 61 AGCGAAGATTGGAGCGCGGAGATTGAGATTCGACAGCAACATGCTGCTGATTT 120
 |||||||
 DB 611 AGCGAAGATTGGAGCGCGGAGATTGAGATTCGACAGCAACATGCTGCTGATTT 670
 OY 121 CACTAGCAAGAGCTGCTCAGGGGAGAAAGCCCTGGCTGCTTAATAATAGTGGGCTGTA 180
 |||||||
 DB 671 CACTAGCAAGAGCTGCTCAGGGGAGAAAGCCCTGGCTGCTTAATAATAGTGGGCTGTA 730
 OY 181 CACACATCAACAGCCAGACAGCGGTGATTGAGACACTCTGTGCCCTGGGGGCTCACT 240

```

|||||
Db 731 CACACATACACCCAGACAGCGGTGTGATGAGACACTGTGCCCCGGGGGCTCAT 790
QY 241 GCCCGTGTCTGCTTGTAACTACTCACTCAGAAATGAGTGTGAGCTGGCTG 300
Db 791 GCCCGTGTCTGCTTGTAACTACTCACTCAGAAATGAGTGTGAGCTGGCTG 850
QY 301 AGCGTGAGTGTGAGTGTCTTGTAAAGGGGAGTGTGAGTGTGAGTGTG 360
Db 851 AGCGTGAGTGTGAGTGTCTTGTAAAGGGGAGTGTGAGTGTGAGTGTG 910
QY 361 TTGACCGCTGTGTGAACATGATGAGTGTGAGTGTGAGTGTGAGTGTG 417
Db 911 TTGACCGCTGTGTGAACATGATGAGTGTGAGTGTGAGTGTGAGTGTG 967

RESULT 6
AF315687 2677 bp mRNA linear PRI 22-MAR-2002
LOCUS AF315687
DEFINITION Homo sapiens S-adenosylhomocysteine hydrolase-like protein mRNA,
ACCESSION AF315687
VERSION AF315687.1 GI:16588686
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2677)
AUTHORS Dekker, J.W., Budhia, S., Angel, N.Z., Cooper, B.J., Clark, G.J.,
Hart, D.N. and Kato, M.
TITLE Identification of an S-adenosylhomocysteine hydrolase-like
transcript induced during dendritic cell differentiation
JOURNAL Immunogenetics 53 (12), 993-1001 (2002)
MEDLINE 21901265
PUBMED 11904675
REFERENCE 2 (bases 1 to 2677)
AUTHORS Dekker, J.W., Budhia, S., Angel, N.Z., Cooper, B.J., Clark, G.J.,
Hart, D.N. and Kato, M.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-2000) Dendritic Cell Research, Mater Medical
Research Institute, Level 3, Aubigny Place, South Brisbane,
Queensland 4101, Australia
FEATURES
Source
Location/Qualifiers
1..2677
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p21"
/clone="211(1)B"
/cell_line="Hodgkin's disease cell line L428"
1..114
/feature="Derived from 5' RACE"
369..1961
/note="PCAL, dendritic cell expressed AHCY-like protein;
expressed in dendritic cells in blood, Langerhans cells
and dermal dendritic cells; mRNA increased during monocyte
differentiation to monocyte-derived dendritic cells in the
presence of GM-CSF and IL-4"
/product="S-adenosylhomocysteine hydrolase-like protein"
/protein_id="AF315687.1"
/db_xref="GI:16588687.1"
/translation="MSMPDAMPPLPGVGEELKQAKIEDEAKTSFMAVTKAPKQIOF
ADDMQETKPTKGRSISQSDSTDSYSSASVTDSDSDSDSDSDSDSDSDSDSDS
SNFCVKNKQAFGRREIEIAEDMSALSLRRAGQERLAGAKIVGCTHITACRAV
LIETICAGACGACMSACNTSYONEVAALAEAGVAFPAKGSSEDDWMCIDRCVNM
DKQAMNILDGCDLTHRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYK
NNDVTKQKFNLYCKRESILDLKRTDVFAGQVYQVYQVYQVYQVYQVYQVYQVYQ
AVYITDIPICALQACMGFVYKINEYIRVDVYITCTGNKNVYTRHLDKRNKSC
IVCMGSHSTEDVTSIRPELTWERVRSQVDHYIWDKRVYTLAEGRLNLSCTV
PFEVSTATTOALALIELYNAPEGRYKQDYVLLPKMDEYVASHLPSFADALTELT
DDQAKYLGINKNGPKPNYRYR"
misc_feature
CDS
#

```

```

BASE COUNT 657 a 648 c 708 g 664 t
ORIGIN
Query Match 100.0%; Score 417; DB 9; Length 2677;
Best Local Similarity 100.0%; Pred. No. 3,7e-117;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGAGAAAGCAGCAACCACTCCAGGGGAGAGCAATTTCTGTGAGAAACATCAAGC 60
Db 643 GAGAGAAAGCAGCAACCACTCCAGGGGAGAGCAATTTCTGTGAGAAACATCAAGC 702
QY 61 AGCAGAAATTTGAGACCGCGGAGATTGAGATTGAGAGCAAGACATGCTGCTGATT 120
Db 703 AGCAGAAATTTGAGACCGCGGAGATTGAGATTGAGAGCAAGACATGCTGCTGATT 762
QY 121 CACTCAGAAAGCTGCTCAGGGGAGAGAGCCCTTGGCTGCTTAAATAGTGGGCTGTA 180
Db 763 CACTCAGAAAGCTGCTCAGGGGAGAGAGCCCTTGGCTGCTTAAATAGTGGGCTGTA 822
QY 181 CACACATACACAGCCAGACAGCGGTGTGATTGAGACACTGTGAGCCCTGGGCTCAGT 240
Db 823 CACACATACACAGCCAGACAGCGGTGTGATTGAGACACTGTGAGCCCTGGGCTCAGT 882
QY 241 GCCCGTGTCTGCTTGTAACTACTCACTCAGAAATGAGTGTGAGTGTGAGTGTG 300
Db 883 GCCCGTGTCTGCTTGTAACTACTCACTCAGAAATGAGTGTGAGTGTGAGTGTG 942
QY 301 AGCGTGAGTGTGAGTGTCTTGTAAAGGGGAGTGTGAGTGTGAGTGTGAGTGTG 360
Db 943 AGCGTGAGTGTGAGTGTCTTGTAAAGGGGAGTGTGAGTGTGAGTGTGAGTGTG 1002
QY 361 TTGACCGCTGTGTGAACATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 417
Db 1003 TTGACCGCTGTGTGAACATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1059

RESULT 7
HSM800298
LOCUS HSM800298 2510 bp mRNA linear PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA; cDNA DKFZP564A1523 (from clone DKFZP564A1523);
ACCESSION AL049954
VERSION AL049954.1 GI:4884203
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2510)
AUTHORS Wambutt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-1999) MIPS, Am Klopferplatz 18a, D-82152
Martinsried, GERMANY
COMMENT clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
sequenced by AGOYA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZP564A1523) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14039
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES
Source
Location/Qualifiers
1..2510
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1, 15.55 cR from D1S248"
/clone="DKFZP564A1523"
/tissue_type="brain"
/clone_id="564 (synonym: hfb22). Vector pAMP1; host
Xl-2blue; sites: NotI + SalI"
/dev_stage="fetal"
1..1796
gene

```

```

CDS
    <1..1796
    /gene="DKFZ564A1523"
    /note="S-adenosyl homocysteine hydrolase homolog XPVKona
    splice variant"
    /product="hypothetical protein"
    /protein_id="CA43223.1"
    /db_xref="GI:4884204"
    /translation="GLITVELDSSVCYILLLRHGRSGRKAIEISQALPLPE
    TOAPSGESENLRRLROGRRAESSONLISRLVSAOGRNGSDRMRLIOLPELDSLE
    ISCFIOIORDMDEFTKEPTKGRSISRSISOSSDYSASSTDSDDVSPRE
    KOOTNRKGSNRCVKNIKOAEGRREIEIEOMSAISIRKRAQGRKPLAGAKITVC
    HTTAQTAVALIETLCALGACQRSACNISTONEVAALAEAVAFAMKSESDDFW
    WCIDRCVNDGQNMNILLDGGDLTWMVKYKKNVEFKIGIYEESTGYHRLYOLSK
    AGRLCYPAHVNDSTKOKEDNLCRCESILDKRTDYMGKGYVQVGVGVG
    CCALKAALGAIYITEIDPICALQACDGEFRRVYKLVNIVQVIVLCTKNKNVTE
    HDLRMNSCTICVMGSHNTEIDTSLRPELTMRVRSOVDHYITMPGKRVLLAER
    LNLSCSTPTPTVLTSTATTQALALIELYNAPGRKQDYTLPLPKMDEVASLHLS
    FNAHLELTDQAKYGLGNKNGPFKPYRY"
    2490

BASE COUNT      664 a      556 c      618 g      672 t
ORIGIN
polya_site
Query Match      99.6%; Score 415.4; DB 9; Length 2510;
Best Local Similarity 99.8%; Pred. No. 1.2e-116;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GAGAGAAGCAGCAACCACTCCAAAGGACACACAAATTTCTGTGAAAGACATCAAGC 60
|||||
478 GAGAGAAGCAGCAACCACTCCAAAGGACACACAAATTTCTGTGAAAGACATCAAGC 537
|||||
61 AGGCAGAAATTTGAGCGCGGAGATTGAGATTGAGAGCAAGACATGTCCTCTGATTT 120
|||||
538 AGGCAGAAATTTGAGCGCGGAGATTGAGATTGAGAGCAAGACATGTCCTCTGATTT 597
|||||
121 CACTCAGAAACGTCCTCAGGGGAGAAAGCCCTGCTGCTGCTAAATAGTGGCTGTA 180
|||||
598 CACTCAGAAACGTCCTCAGGGGAGAAAGCCCTGCTGCTGCTAAATAGTGGCTGTA 657
|||||
181 CACACATCAGAGCCCAAGCAGCGGTGATTGAGACACTGTGCGCTGGGGGCTCAGT 240
|||||
658 CACACATCAGAGCCCAAGCAGCGGTGATTGAGACACTGTGCGCTGGGGGCTCAGT 717
|||||
241 GCCCGTGTCTGCTTGTAACTCACTCACTCAAGATGAAGTCACTCAGACATGGCTG 300
|||||
718 GCCCGTGTCTGCTTGTAACTCACTCACTCAAGATGAAGTCACTCAGACATGGCTG 777
|||||
61 AGGCAGAAATTTGAGCGCGGAGATTGAGATTGAGAGCAAGACATGTCCTCTGATTT 360
|||||
778 AGGCAGAAATTTGAGCGCGGAGATTGAGATTGAGAGCAAGACATGTCCTCTGATTT 837
|||||
361 TTGACCGCTGTGTAACATGATGGGTGGGAGGCAACATGATCCTGTGATGAGGG 417
|||||
838 TTGACCGCTGTGTAACATGATGGGTGGGAGGCAACATGATCCTGTGATGAGGG 894
|||||

RESULT 8
BC018218      1772 bp      mRNA      linear      ROD 07-AUG-2002
LOCUS
DEFINITION
MUS MUSCULUS, S-adenosylhomocysteine hydrolase-like 1, clone
BC018218
IMAGE:4007102, mRNA, complete cds.
ACCESSION
BC018218.1 GI:17390492
VERSION
MGC.
KEYWORDS
house mouse.
SOURCE
MUS MUSCULUS
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1772)
AUTHORS
Strausberg R.
TITLE
Direct Submission
JOURNAL
Submitted (03-DEC-2001) National Institutes of Health, Mammalian

```

```

REMARK
COMMENT
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NHI-MGC Project URL: http://ngc.ncl.nih.gov
Contact: MGC help desk
Email: gcgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgdbcm.tmc.edu
Guanarone, P.H., Garcia, A.M., Lu, X., Huik, S.W., Hale, S.M.,
Yoon, V.S., Kowls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov.
Series: IRAX Plate: 23 Row: n Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency OFE
analysis, Genomescan gene prediction, similarity but not identity
to protein.

FEATURES
source
1..1772
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="CECH II"
/clone="MGC:18748 IMAGE:4007102"
/tissue_type="Mammary tumor metastasized to lung,
MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR
enhancer."
/clone_id="NCL CGAP Lu30"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
191..1693
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH18218.1"
/db_xref="GI:17390493"
/db_xref="locusID:10768"
/translation="MATVTKAPKQIOFADMOEFTKPTTGRSLHSISOSTDS
YSSASYSDDSDDESPREKQOTNSGNSNFCVKNIKOAEGRREIEIEADMSALIS
LKRROGEPPLAGAKIVGCTHTTAQTAVALIETLCALGACQRSACNISTONEVAAL
AEAVAFAMKSESDDFWMCIDRCVNDGQNMNILLDGGDLTWMVKYKKNVEFKKI
RGIVSESYGYHRLYOLSKAGRLCYPAHVNDSTKOKEDNLCRCESILDKRTDY
VMFGKQYVGVGEGYVGGCCCAALKAALGAIYITEIDPICALQACDGEFRRVYKLVN
I RQDVAVITCTGNKNVYTRHDLRKNNSCTICVMGSHNTEIDTSLRPELTMRVRSO
VDHYITMPGKRVLLAERLNLSCSTPTPTVLTSTATTQALALIELYNAPGRKQD
VYLLPKMDEVASLHLSFNAHLELTDQAKYGLGNKNGPFKPYRY"

BASE COUNT      462 a      406 c      506 g      398 t
ORIGIN
Query Match      92.3%; Score 385; DB 10; Length 1772;
Best Local Similarity 95.2%; Pred. No. 2.6e-107;
Matches 397; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

1 GAGAGAAGCAGCAACCACTCCAAAGGACACACAAATTTCTGTGAAAGACATCAAGC 60
|||||
375 GAGAGAAGCAGCAACCACTCCAAAGGACACACAAATTTCTGTGAAAGACATCAAGC 434
|||||
61 AGGCAGAAATTTGAGCGCGGAGATTGAGATTGAGAGCAAGACATGTCCTCTGATTT 120
|||||
435 AGGCAGAAATTTGAGCGCGGAGATTGAGATTGAGAGCAAGACATGTCCTCTGATTT 494
|||||
121 CACTCAGAAACGTCCTCAGGGGAGAAAGCCCTGCTGCTGCTAAATAGTGGCTGTA 180
|||||
495 CACTCAGAAACGTCCTCAGGGGAGAAAGCCCTGCTGCTGCTAAATAGTGGCTGTA 554
|||||
181 CACACATCAGAGCCCAAGCAGCGGTGATTGAGACACTGTGCGCTGGGGGCTCAGT 240
|||||

```

Db 555 CACACATACGCGCCGACAGCGGTATTAATTGAGACCGCTTGTGCGCCGAGAGCTCACT 614
 QY 241 GCCGCTGCTGCTGTGTAACATCTCACTCAGAAATGAGTAGCTGACAGACTGGCTG 300
 Db 615 GCCGCTGCTGCTGTAACATCTCACTCAGAAATGAGTAGCTGACAGACTGGCTG 674
 QY 301 AGCGTGGAGTTGACATGCTTTCCTTGAAGGGCGAGTCAAGATGACTTCTGGTGTGA 360
 Db 675 AGCGTGGAGTTGACATGCTTTCCTTGAAGGGCGAGTCAAGATGACTTCTGGTGTGA 734
 QY 361 TTGACCGCTGTGTGAACATGATGGTGGCAGCGCAACATGATCCGATGATGGGG 417
 Db 735 TTGACCGCTGTGTGAACATGATGGTGGCAGCGCTAATGATCTCGATGATGGGG 791

RESULT 9
 AK025372 1987 bp mRNA linear PRI 29-SEP-2000
 LOCUS Homo sapiens CDNA: FLJ21719 f1s, clone COLF0094.
 DEFINITION AK025372.1 GI:10437875
 VERSION AK025372.1
 KEYWORDS oligo capping; f1s (full insert sequence).
 SOURCE Homo sapiens colon mucosa CDNA to mRNA, clone_11b:COLF clone:COLF0094.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Ota, T., Suzuki, Y., Ota, T., Oabayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
 TITLE Direct Submission
 JOURNAL NEDO human CDNA sequencing project
 REFERENCE 2 (bases 1 to 1987)
 AUTHORS Sugano, S., Suzuki, Y., Ota, T., Oabayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
 TITLE Direct Submission
 JOURNAL NEDO human CDNA sequencing project supported by Ministry of International Trade and Industry of Japan; CDNA full insert sequencing: Research Association for Biotechnology; CDNA library construction, 5' - 3' - end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by science and Technology Agency).

FEATURES
 SOURCE 1. 1987
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="COLF0094"
 /tissue_type="colon mucosa"
 /clone_id="COLF"
 /note="cloning vector pME18SFL3"
 BASE COUNT 557 a 430 c 501 g 499 t
 ORIGIN
 Query Match: 66.2%; Score 276.2; DB 9; Length 1987;
 Best Local Similarity 78.9%; Pred. No. 8e-74;
 Matches 329; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1 GAGGAGACACCAACCACTCCAGGAGAGAGCAATTCCTGTGAGAGCAATCAAC 60
 Db 480 GGGAGACACCAACCACTCCAGGAGAGAGCAATTCCTGTGAGAGCAATCAAC 539
 QY 61 AGGAGATTTGAGAGCGCGGAGATTGAGATGACAGACAGATGCTGCTGATTT 120
 Db 540 AGGAGATTTGAGAGCGCGGAGATTGAGATGACAGACAGATGCTGCTGATTT 599
 QY 121 CACTCAGGAAACGTGCTCAGGGGGAGAACCCCTTGCTGCTAAATATGCTGCTGA 180

Db 600 CTTTGGAGAGAGAGCTCAAGAGAGAAAGCCCTTGGCTGAGGCCAAATGCTGGGTGCA 659
 QY 181 CACACATACAGGCCAGACAGCGGTGTGATGATGACACATCTGCGCCGTGGGCTCAGT 240
 Db 660 CACACATACAGGCCAGACAGCGGTGTGATGATGACACATCTGCGCCGTGGGCTCAGT 719
 QY 241 GCCGCTGCTGCTGTGTAACATCTCACTCAGAAATGAGTAGCTGACAGACTGGCTG 300
 Db 720 GCCGCTGCTGCTGTGTAACATCTCACTCAGAAATGAGTAGCTGACAGACTGGCTG 779
 QY 301 AGCGTGGAGTTGACATGCTTTCCTTGAAGGGCGAGTCAAGATGACTTCTGGTGTGA 360
 Db 780 AGCGTGGAGTTGACATGCTTTCCTTGAAGGGCGAGTCAAGATGACTTCTGGTGTGA 839
 QY 361 TTGACCGCTGTGTGAACATGATGGTGGCAGCGCAACATGATCCGATGATGGGG 417
 Db 840 TTGACCGCTGTGTGAACATGATGGTGGCAGCGCTAATGATCTCGATGATGGGG 896

RESULT 10
 BC008349 2030 bp mRNA linear PRI 12-JUL-2001
 LOCUS Homo sapiens, similar to S-adenosylhomocysteine hydrolase-like 1, clone IMAGE:3536052, mRNA, partial cds.
 DEFINITION BC008349.1 GI:14249935
 VERSION BC008349.1
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 2030)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: RbIn Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guln, Leticia Hsiao, Martin Krzyzanski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 23 Row: 5 Column: 18.
 Location/Qualifiers
 1. 2030
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3536052"
 /tissue_type="lung, small cell carcinoma"
 /clone_id="NIH-MGC-7"
 /lab_host="DH10B-R"
 /note="Vector: pORF7"
 <1..1838
 /codon_start=3
 /product="Similar to S-adenosylhomocysteine hydrolase-like 1"

FEATURES
 source
 CDS

1
 /product="Similar to S-adenosylhomocysteine hydrolase-like 1"

VERSION BC024325.1 GI:19353091
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 5052)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgaps-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mdickpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 22 Row: 1 Column: 22
 This clone was selected for full length sequencing because it passed the following selection criteria: similarity but not identity to protein.
 Location/Qualifiers
 1. 5052
 /organism="Homo sapiens"
 /db_xref="locusID:23382"
 /db_xref="taxon:9606"
 /clone="MGC:21525 IMAGE:3907552"
 /issue_type="uterus, leiomyosarcoma"
 /clone_id="NH_MGC_71"
 /lab_host="DH10b"
 /note="Vector: pCMV-SPORT6"
 55..1890
 /codon_start=1
 /product="KIAA0828 protein"
 /protein_id="AAH24325.1"
 /db_xref="GI:19353092"
 /translation="MSYQVYSAAASAAARVPEVLEKDISPSEASOLGISTAVGAMAP AGGDEPAPAPAEPRYPVPGSGPAAALSPACKVPOASAMKRSDEPHHQRHDDG EALVSPGVTYEAAPRYVKQIQFADQKQENKRPRTKIGRSLRSISOSTDYSSAA SYTDSDEPSPDKQOKNSGSDPCVKNKQAEFRREIEIEOEWPALMAKRA OGEKPLAGAKIVGCTHTTAQTAVMETGALAGCQMAACNITSTLMEVPAALAEGR PFAFMGSESDPMMICIDRCVNEGMOPTNDGDLTHIYKRYKPMFAKIGIVE ESYTGHARLQLSKAKRLCYPRANNVNDYVKOKPDNACCSHLDLQIKRTDMKEG KQVAVCGTGEVGGCCALAMGSIYVTEIDPICALQACDSGRVLKLVNIVQVDI VITCGKRNVTREHDLRMKNSCIVCMNGSNSEIDVASLTPELWERYVSQVDHI WPGKRIVLLAEGRLNLSCSTVPTFVLSITATQALALIELVNAPEERYKQDYLIP KMDVEYASHLPTEDAHLELDEQAKYGLNKGFPKRYNYR"
 BASE COUNT 1300 a 1184 c 1145 g 1423 t
 ORIGIN
 Query Match 65.9%; Score 274.6; DB 9; Length 5052;
 Best Local Similarity 78.7%; Pred. No. 2.5e-73;
 Matches 328; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
 Oy 1 GAGGAGAACGACCAACCACTCCAGGACGACGACATTTCTGTGTAAGAACATCAAC 60
 Db 572 GGGGACAGCAGCAAAAGACCTTAAGGAGACAGGACTTCTGTGTAAGAACATCAAC 631
 Oy 61 AGGAGAAATTGGAGCGCGGAGATGAGATTGAGAGACAGACATGCTGCTGATTT 120
 Db 632 AGGAGAAATTGGAGCGAGAAAGAAATTTGTAATTCGTAACAGAAATGCGCTGATGATGG 691

Oy 121 CACGAGGAAGACGCTGCTCAGGGGAGAAAGCCCTTGCGTGTCTAAATAGTGGGCTGA 180
 Db 692 CTTTGAAGAGAGACGCTCAAGAGAAAGCCTTTGGCTGGACCAAAATCGGGTTGA 751
 Oy 181 CACACATCAGACCCAGACAGCGGTGTTGATGAGACATCTGTGCGCCGTGGGGGCTCACT 240
 Db 752 CACACATCAGCTGCTCAGACGCTGCTGTCTTATGGAACCTGTGGGTGCTCGGGGCCCACT 811
 Oy 241 GCGGCTGCTGCTGTTGTAACATCTACTCACTCAGATGAGTACAGTGCAGCATGCGCTG 300
 Db 812 GCGGATGGGCTGCTGCAACATCTATTCCTCCTCAATGAGTGGCTGCTGCTAGCAG 871
 Oy 301 AGGCTGAGTTCAGTGTGCTGTTGAGGAGGAGCAGCAAGATGACTCTGTGGTGTGA 360
 Db 872 AAGTGATTTCTGCTTTTGGCTGGAGAGAGTCAGAAAGTACCTTTGGTGGTGA 931
 Oy 361 TTGACCGCTGTGTAACATGATGAGTGGTGGCAGCCCAACATGATCTGTGATGATGGG 417
 Db 932 TCGATAGATGTGTGATGTGAGGCGGTGCGACGCAACATGATCTGTGATGATGGAG 988

RESULT 13
 AY113501
 LOCUS
 DEFINITION Drosophila melanogaster R58316 full insert cDNA.
 ACCESSION AY113501
 VERSION AY113501.1 GI:21064552
 KEYWORDS FLI.CDNA.
 SOURCE Drosophila melanogaster.
 ORGANISM Drosophila melanogaster.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1792)
 Stapleton, M., Brockstein, P., Hong, L., Agbayani, A., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G., Miranda, A., Mungall, C.J., Nunoo, J., Pacleb, J., Paragias, V., Park, S., Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Rubin, G.M., and Celisner, S.
 Direct Submission
 Submitted (16-MAY-2002) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA
 Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory
 Berkeley, CA 94720

TITLE
 JOURNAL
 COMMENT
 This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our web site (<http://fruitfly.berkeley.edu>) or send email to cdna@fruitfly.berkeley.edu.

FEATURES
 source
 1..1792
 /organism="Drosophila melanogaster"
 /strain="y; cn bw sp"
 /db_xref="taxon:7227"
 /map="89E10-89E10"
 1..1792
 /gene="Abcy89e"
 /note="alignment with genomic scaffold AE003715"
 /db_xref="FlyBase:FBgn0015011"

Query Match	45.8%	Score 191;	DB 3;	Length 1792;
Best Local Similarity	66.3%	Pred. No.	1.3e-47;	
Matches 275;	Conservative	0;	Mismatches 140;	Indels 0;
				Gaps 0;

RESULT_14			
AC020395/c			
LOCUS			
DEFINITION			
AC020395	52423 bp	DNA	linear
Drosophila melanogaster, ***	SEQUENCING IN PROGRESS	***,	in ordered

FEATURES	source	1. 52423	/organism="Drosophila melanogaster"
BASE COUNT	15554 a 10893 c 10667 g 15309 t		
ORIGIN			
Query Match	45.8%; Score 191; Db 2; Length 52423;		
Best Local Similarity	66.3%; Pred. No. 1.4e-47;		
Matches 275; Conservative 0; Mismatches 140; Indels 0; Gaps 0;			
* NOTE: This is a 'working draft' sequence.			
* This sequence will be replaced			
* by the finished sequence as soon as it is available and			
* the accession number will be preserved.			
Location/Qualifiers			
3	GAGAAAGCAGCAACACATTCCTCCAGGGGCGAGCAACATTTCTGTGTGAATAACATCAAGCAG 62		
Db 43851	GACAAACCATTCAGCCGCACTCCGCGGAGGCACTGACTTTTSCGTGMAAGATTCGCAAG 43792		
QY 63	GCAGAAATTTGGACGCCGCGGAGATTGAGATTTCAGAGCAAGACATTCGTCTGTGATTCA 122		
Db 43791	AGTGATTTGGAAAGCGGGAGATGAGATGATCCCGAGTGGGAGATGCCGGGCAATCATGACT 43732		
QY 123	CTCAGGAAACGTGCTCAGAGGGGAGAAAGCCCTTGGCTGTCTAAATATGTGGGCTGTACA 182		
Db 43731	CTGAGGAAAGAGGAGAAAGATGAGAGAGCCCTTAAGAGGTGGCCAAATATGTCGGAGACCC 43672		
QY 183	CACATCAACAGCCGAGACAGCGGTGTGATTAGACATCTGTGCCCTGGGGGCTCAGTGC 242		
Db 43671	CACGTCATGCTCAGTGGCGAGTGTGTATTCAGACCCCTGTCCAACTGTGGGGCCCACTT 43612		
QY 243	CGCTGTGCTGCTTTGATCATCTACTCAACTCAGATGAATGAATGATGCTGACAGCACTGGCTGAG 302		
Db 43611	CGCTGGGCTGCTCCACAAATTATTCACAAAGAGATGTCGGCGCTCTGGCAGAG 43552		
QY 303	GCTGGAGTTGAGATGTTGGCTTGGAAAGGGGAGATCAGAAAGATGACTTGTGGTGGTATT 362		
Db 43551	GGCGGAATTCGCACTTCTGCTCGCGCGGAGAGACGAGGAGAGACTTGTGTGTGCTTG 43492		
QY 363	GACCGCTGTGTGAACATGATGATGGTGGCAGGCGCAACATGATCTGTGATGATGGGG 417		
Db 43491	GACAGGGCCATCTACTCCGACGCGTGGCAGCGCAACCTAATCTGTGACGACGCGCG 43437		
RESULT 15			
DROABDB			
LOCUS	DROABDB 80423 bp DNA linear INV 13-OCT-1995		
DEFINITION	Drosophila melanogaster (clone DS07696) Abd-B gene.		
ACCESSION	L07835		
VERSION	L07835.1 GI:156737		
KEYWORDS			
SOURCE	Drosophila melanogaster (tissue library: P1) adult whole fly DNA.		
ORGANISM	Drosophila melanogaster		
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (sites)		
REFERENCE	Regulski, M., Harding, K., Kostriken, R., Karch, F., Levine, M. and McGinnis, W.		
TITLE	Homeo box genes of the Antennapedia and bithorax complexes of Drosophila		
JOURNAL	Cell 43 (1), 71-80 (1985)		
MEDLINE	86079516		
REFERENCE	2416463		
REFERENCE	2 (sites)		
REFERENCE	DeLongenzi, M., Ali, N., Saari, G., Henry, C., Wilcox, M. and Blenz, M.		
TITLE	Evidence that the Abdominal-B r element function is conferred by a trans-regulatory homeoprotein		
JOURNAL	EMBO J. 7 (10), 3223-3231 (1988)		
MEDLINE	89030618		
REFERENCE	2903049		
REFERENCE	3 (sites)		
REFERENCE	Zavortink, M. and Sakonju, S.		

TITLE

The morphogenetic and regulatory functions of the Drosophila Abdominal-B gene are encoded in overlapping RNAs transcribed from separate promoters

JOURNAL
Genes Dev. 3 (12a), 1969-1981 (1989)

MEDLINE
90152341

PUBMED
2482824

REFERENCE
4. (bases 1 to 80423)

AUTHORS
Martin,C.H., Gelniker,S.E., Davis,C.A., Mayeda,C.A.,
Stratmann,M.P., Yoshida,K. and Palazolo,M.J.

JOURNAL
Unpublished (1992)

FEATURES

Location/Qualifiers
1..80423
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/map="3, 89E; Polytene"
/tissue_type="whole fly"
/dev_stage="adult"
/tissue_lld="p1"

misc_feature 1..80423
BASE COUNT 23439 a 16747 c 16641 g 23596 t
ORIGIN

Query Match

45.8%; Score 191; DB 3; Length 80423;

Best Local Similarity 66.3%; Pred. No. 1.5e-47;

Matches 275; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 3 GAGAGCAGCAAACTCCAGAGGAGCAACATTTCTGTGTAAGAACATCAAGCAG 62
DB 3342 GACACACATCAGCAGCACTCCGCGAGGCACTGACTTTGGGTGAAGACATCTCGAAG 3401
QY 63 GCAGATTGGAGCGCGGAGATGAGATGAGAGCAAGACATCTCTCTGATTCA 122
DB 3402 AGTGATTGGAGGCGGAGATCGAGATCGGAGATCGCGGAGATCGCGGAGATCGACT 3461
QY 123 CTCAGAAACGCTCAGAGGAGAGAGCCCTGGCTGCTAAATAGTGGGCTGACA 182
DB 3462 CTGAGAAAGAGCAGAGATGAGAGCCCTAAAGGTGCAATATCTCGATGACCC 3521
QY 183 CACATCAGACCCAGACAGCGGCTGTGATGAGACACTCTGCCCCGGGGGCTCAGTGC 242
DB 3522 CACGTCATGCTCAGTCGAGTGTGATGAGAGACCCCTGTCACACTGGGCGCACAGTT 3581
QY 243 CGCTGTCTGCTTAAATCATCTACCACTCAATGAATGAATGAGTGTGAGTGTGAG 302
DB 3582 CGCTGGGCTGCTGCAACATTTATTCACACAAAGCAGTGTGCGCCCTCTGGCAGAG 3641
QY 303 GCTGAGTTGAGTGTGCTGAGAGGCGAGTCAAGATGACTTCTGTGGTGTATT 362
DB 3642 GCGGAAATTCGATCTTGCTGCGCGGAGAGAGAGAGAGTCTGTGTGTCTTG 3701
QY 363 GACCGCTGTGATGATGATGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 417
DB 3702 GACAGGGCCATCTACTCGAGCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3756

Search completed: April 21, 2003, 02:39:04
Job time : 1107.44 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 23:48:04 ; Search time 18.6401 Seconds

(without alignments)
6860.687 Million cell updates/sec

Title: US-09-782-051-1_COPY_529_945
 Page: 417

Sequence: 1 gagagaagcagcaaaccaac.....catgatcctgtagtggg 417

Scoring table: IDENTITY_NUC

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 08

Listing first 45 summaries

Database : Issued_Patents_NA:*

```

1: /cgr2_6/prodata/1/ma/5A_COMB.seq: *
2: /cgr2_6/prodata/1/ma/5B_COMB.seq: *
3: /cgr2_6/prodata/1/ma/6A_COMB.seq: *
4: /cgr2_6/prodata/1/ma/6B_COMB.seq: *
5: /cgr2_6/prodata/1/ma/PCTTS_COMB.seq: *
6: /cgr2_6/prodata/1/ma/PCTTS_Less.seq: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Length	DB	ID	Description
C	1	226.4	54.3	636	4	US-09-328-111-850	Sequence 850, Apple
C	2	138.8	33.3	2211	4	US-09-318-448-26	Sequence 26, Apple
C	3	138.8	33.3	2211	4	US-09-347-878-2	Sequence 2, Apple
C	4	118.2	28.3	1812	2	US-08-669-536-1	Sequence 1, Apple
C	5	101.0	24.4	1767	3	US-08-930-894-1	Sequence 1, Apple
C	6	72.6	17.4	289	1	US-08-204-140-8	Sequence 8, Apple
C	7	72.6	17.4	289	3	US-09-081-1674-8	Sequence 8, Apple
C	8	72.6	17.4	289	3	US-09-081-395-8	Sequence 8, Apple
C	9	72.6	17.4	289	4	US-09-146-833-8	Sequence 8, Apple
C	10	72.6	17.4	289	5	PCT-US95-02521-8	Sequence 8, Apple
C	11	68.4	16.4	285	1	US-08-204-740-6	Sequence 6, Apple
C	12	68.4	16.4	285	3	US-09-081-1674-6	Sequence 6, Apple
C	13	68.4	16.4	285	3	US-09-081-395-6	Sequence 6, Apple
C	14	68.4	16.4	285	4	US-09-416-833-6	Sequence 6, Apple
C	15	68.4	16.4	285	5	PCT-US95-02521-6	Sequence 6, Apple
C	16	33.6	8.1	7218	1	US-08-332-463-14	Sequence 14, Apple
C	17	32.8	7.9	31571	1	US-08-323-443B-1	Sequence 1, Apple
C	18	32.8	7.9	53526	3	US-08-658-136-2	Sequence 2, Apple
C	19	32.8	7.9	53577	3	US-08-658-136-1	Sequence 1, Apple
C	20	31.2	7.5	233	2	US-08-687-080-108	Sequence 108, Apple
C	21	31.1	7.4	3306	4	US-09-170-170-7	Sequence 7, Apple
C	22	30.2	7.2	1317	4	US-09-134-078-2	Sequence 2, Apple
C	23	29.4	7.1	4483	4	US-08-561-527-363	Sequence 363, Apple
C	24	29.4	7.1	12394	4	US-09-488-856-10	Sequence 10, Apple
C	25	29	7.0	1803	1	US-08-021-1080-7	Sequence 7, Apple
C	26	29	7.0	1803	1	US-08-726-160-7	Sequence 7, Apple
C	27	29	7.0	1803	5	PCT-US94-01783-7	Sequence 7, Apple

ALIGNMENTS

C 28	29	7.0	2381	1	US-08-021-608D-9	Sequence 9, Appl 1
C 29	29	7.0	2381	1	US-08-021-608D-9	Sequence 9, Appl 1
C 30	29	7.0	2381	1	US-08-726-160-9	Sequence 9, Appl 1
C 31	29	7.0	2381	5	PCR-US94-01782-1	Sequence 9, Appl 1
C 32	29	7.0	2384	1	US-08-021-608D-9	Sequence 1, Appl 1
C 33	29	7.0	2384	1	US-08-726-160-1	Sequence 1, Appl 1
C 34	28.8	6.9	2384	5	PCR-US94-01782-1	Sequence 1, Appl 1
C 35	28.6	6.9	18627	4	US-08-961-527-113	Sequence 113, Appl 1
C 36	28.6	6.9	3744	4	US-08-961-527-263	Sequence 263, Appl 1
C 37	28.2	6.8	5816	4	US-09-020-641-4	Sequence 4, Appl 1
C 38	28.2	6.8	3879	4	US-08-916-352-1	Sequence 1, Appl 1
C 39	28.2	6.8	3979	4	US-09-085-1998B-10	Sequence 10, Appl 1
C 40	28	6.7	6827	4	US-08-961-527-60	Sequence 60, Appl 1
C 41	28	6.7	308	1	US-08-627-706-5	Sequence 5, Appl 1
C 42	28	6.7	308	4	US-09-103-489-5	Sequence 5, Appl 1
C 43	28	6.7	500	1	US-08-627-706-9	Sequence 9, Appl 1
C 44	28	6.7	500	4	US-09-103-489-9	Sequence 9, Appl 1
C 45	27.8	6.7	1875	4	US-09-761-716-1	Sequence 1, Appl 1
			1611	2	US-08-650-596-1	Sequence 1, Appl 1

RESULT 1
US-09-328-111-850/c

; Sequence 850, Application US/09328111
; Patent No. 6262333

APPLICANT: Endege,

APPLICANT: Astle, Jon H.

APPLICANT: Bushnell, Steven E.

APPLICANT: Catlino, Theodore J.

APPLICANT: Ford, Donna M

APPLICANT: Monahan, John E.

; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

; TITLE OF INVENTION: FRODOCS
; FILE REFERENCE: CCD-257 (US)

CURRENT FILING DATE: 1999-06-08

EARLIER FILING DATE: 1998-06-10

```

; NUMBER OF SEQ ID NOS: 830
; SOFTWARE: FastSeq for Windows Version 3.0

```

; SEQ AD NO 6
; LENGTH: 6

```

; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; FEATURE:
NAME/KEY: misc_feature
;

```

```

; LOCATION: (1):::(0508)
; OTHER INFORMATION: n = A,T,C or G

```

0509-328-111-850

Query Match
Best Local

Matches 23

QY - 1/9 TAC III

420 JAC

QY 238 AGT 111

DB 360 AGT

QY 298 CIG III

DB 300 CTGAGGCTGGAGTTCAGATGCTTGGAGGCGAGTCAAGATGACTTGTGTGT 241
 QY 358 GTATTGACCGCTGCTGCAATGATGCGGAGCCCAACATGCTGATGATGGG 417
 DB 240 GTATTGACCGCTGCTGCAATGATGCGGAGCCCAACATGATGCTGATGATGGG 181

RESULT 2

US-09-318-448-26

Sequence 26, Application US/09318448
 Patent No. 6210950
 GENERAL INFORMATION:
 APPLICANT: Johnson, William G.
 APPLICANT: Steenroos, Edward S.
 TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
 TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
 FILE REFERENCE: 601-1-057
 CURRENT APPLICATION NUMBER: US/09/318,448
 CURRENT FILING DATE: 1999-05-25
 NUMBER OF SEQ ID NOS: 46
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 26
 LENGTH: 2211
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-318-448-26

Query Match 33.3%; Score 138.8; DB 4; Length 2211;
 Best Local Similarity 61.2%; Pred. No. 5.7e-36;
 Matches 224; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 52 ACATCAAGCAGGAGATTTGGAGCGCGGAGATTGAGATTCAGACAGATGCTG 111
 DB 79 ACATGCGCTGCTGCTGCTGCGGAGCGAGGCGCTGACATTCGTGAGAGAGATGCGG 138
 QY 112 CTCGTATTCTACAGAGAACTGCTCAGGAGGAGAGAGCCCTGGCTGCTGCTAAATAG 171
 DB 139 GCCTGATGCTATGCGGAGCGGAGTACTGCGCTCCAGGCACTGAAGGGGCGCGATCG 198
 QY 172 TGGGCTGTACACATACACAGCCAGACGCGGTGTGATGAGACACTGTGCTGCTG 231
 DB 199 CTGGCTGCTGCAACATGACCGCGGAGCGGCGCTGCTGATGAGACCTGCTCAGCTG 258
 QY 232 GGGCTGAGTCCGCTGCTGCTGCTGATCACTCACTCACTCACTCACTCACTCACT 291
 DB 259 GTCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 318
 QY 292 CACTGCTGAGGCTGAGATTCAGATTCCTGCTGAGAGGCGAGTCAAGATGACTTCT 351
 DB 319 CCATTCGCAAGGCTGCGATTCGCGTGTATGCTGGAAGGGGAGAAAGAGAGATACC 378
 QY 352 GGTGCTGTATGACCGCTGTGTGAACATGAGGAGGAGGCGCCACATGATCTGTGATG 411
 DB 379 TGTGCTGATGAGACAGCCCTGTACTTCAAGAGCGGCGCCCTCAACATGATCTGTGACG 438
 QY 412 ATGGGG 417
 DB 439 ACGGGG 444

RESULT 3

US-09-347-878-2

Sequence 2, Application US/09347878C
 Patent No. 6376210
 GENERAL INFORMATION:
 APPLICANT: Yuan, Chong
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
 FILE REFERENCE: 25885-1651
 CURRENT APPLICATION NUMBER: US/09/347,878C
 CURRENT FILING DATE: 1999-07-06
 NUMBER OF SEQ ID NOS: 75
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 2

LENGTH: 2211
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: Human S-adenosylhomocysteine hydrolase cDNA
 PUBLICATION INFORMATION:
 DATABASE ACCESSION NUMBER: M61831/GenBank
 US-09-347-878-2

Query Match 33.3%; Score 138.8; DB 4; Length 2211;
 Best Local Similarity 61.2%; Pred. No. 5.7e-36;
 Matches 224; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 52 ACATCAAGCAGGAGATTTGGAGCGCGGAGATTGAGATTCAGACAGATGCTG 111
 DB 79 ACATGCGCTGCTGCTGCTGCGGAGCGAGGCGCTGACATTCGTGAGAGAGATGCGG 138
 QY 112 CTCGTATTCTACAGAGAACTGCTCAGGAGGAGAGAGCCCTGGCTGCTGCTAAATAG 171
 DB 139 GCCTGATGCTATGCGGAGCGGAGTACTGCGCTCCAGGCACTGAAGGGGCGCGATCG 198
 QY 172 TGGGCTGTACACATACACAGCCAGACGCGGTGTGATGAGACACTGTGCTGCTG 231
 DB 199 CTGGCTGCTGCAACATGACCGCGGAGCGGCGCTGCTGATGAGACCTGCTCAGCTG 258
 QY 232 GGGCTGAGTCCGCTGCTGCTGCTGATCACTCACTCACTCACTCACTCACTCACT 291
 DB 259 GTCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 318
 QY 292 CACTGCTGAGGCTGAGATTCAGATTCCTGCTGAGAGGCGAGTCAAGATGACTTCT 351
 DB 319 CCATTCGCAAGGCTGCGATTCGCGTGTATGCTGGAAGGGGAGAAAGAGAGATACC 378
 QY 352 GGTGCTGTATGACCGCTGTGTGAACATGATGAGGAGGCGCCACATGATCTGTGATG 411
 DB 379 TGTGCTGATGAGACAGCCCTGTACTTCAAGAGCGGCGCCCTCAACATGATCTGTGACG 438
 QY 412 ATGGGG 417
 DB 439 ACGGGG 444

RESULT 4

US-08-669-536-1

Sequence 1, Application US/08669536
 Patent No. 5910444
 GENERAL INFORMATION:
 APPLICANT: MASUTA, CHIKARA
 APPLICANT: UENHARA, KYOKO
 APPLICANT: TANAKA, HIDEO
 APPLICANT: KUNIMATA, SHIGERU
 TITLE OF INVENTION: ORGANISMS IN WHICH THE EXPRESSION OF
 TITLE OF INVENTION: S-ADENOSYLHOMOCYSTEINE HYDROLASE GENE IS INHIBITED
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
 STREET: PO BOX 747
 CITY: FALLS CHURCH
 STATE: VA
 COUNTRY: USA
 ZIP: 22040-0747
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/669,536
 FILING DATE:
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: MURPHY JR, GERALD M
 REGISTRATION NUMBER: 28,977

TELEPHONE/CODER NUMBER: 1254-148
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1812 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
DS-08-666-536-1

Query Match	28.3%;	Score 118.2;	DB 2;	Length 1812;
Best Local Similarity	57.6%;	Pred. No. 3.2e-29;		
Matches 232; Conservative	0;	Mismatches 168;	Indels 3;	Gaps 1

QY	18	AACCTCAAGGGCAGACAGCAATTTCTGTGTGAAAGCAACATCAAGCGAGAAATTTGAGCGC	77
Db	82	AAGACACACCTCTGGCCCGGAGATACAAAGTCAAGACATGTCTCAGGCGCAATTCGGCCGG	144
QY	78	CGGAGATTGAGATTGCAGACCAAGACATGTCTGTCTGATTTTCATCAGAGAAACGTCT	137
Db	142	CTTGAAATTCGAGTGTGGCCGAAAGTTGAAATGCGCTGTCTCATGGCTTGTGACTGAATTT	201
QY	138	CAGGGGGAGAAAGCCCTTGGCTGTGGTGTCTTAAATTAATGGGCGTGTACACACATCACAGCCAG	197
Db	202	GGCCCTTCACAGCCATTTAAAGGTGCTAAAGTTACTGGATCTTTACATATGACCATTTCAA	261
QY	198	ACACGCGGTGATTGTGAGACACTGTGTCCCTGTGGGGCTCAGTCCGCTGGTCTGCTTGT	257
Db	262	ACTGCAAGTTTGTGATTGAAACCCCTTACTGTCTTGGGTGTCTGAAGTTAATGTGTCTTTCG	322
QY	258	AACATCTACTCACTCAAGTGAAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	317
Db	322	AACATCTCTCTCACTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	381
QY	318	TTCCCTTGGAAAGGGCGAGTCAAGAAATGACTTCTGGTGTGTATATTGACCGCTGTGTGAAC	377
Db	382	TTCCGCTGGAAAGGGGTGAGACTCTCAGAGAGATATGGTGTGTACTGTGAGAGGGCACTTGAC	441
QY	378	ATGAT---GGGTGGCAGAGCCCAACATGATCTCTGGATGTGGGG	417
Db	442	TGGGTCCAGGTGTGTGGGCCCGAATTGTATCTCTCAACATGTGTG	484

RESULT 5
US-08-930-894-1

GENERAL INFORMATION:
APPLICANT: GREENLAND, Andrew James
APPLICANT: DRAPER, John
APPLICANT: SKIPSEY, Marc
APPLICANT: WARNER, Simon
TITLE OF INVENTION: 5-ADENOSYL-L-HOMOCYSTEIN HYDROLASE PROMOTER
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pillsbury Madison & Sutro
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/930,894
FILING DATE: 09-OCT-1997
CLASSIFICATION: 800

```

PRIORITY APPLICATION DATA
APPLICATION NUMBER: PCT/GB96/00882
FILING DATE: 10-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9507381.3
FILING DATE: 10-APR-1995
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1767 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: SHH GENE FROM ASPARAGUS
FEATURE:
NAME/KEY: CDS
LOCATION: 26..1483
OTHER INFORMATION: /codon_start= 2
US-08-930-894-1

```

Query Match	24.48;	Score 101.8;	DB 3	Length 1767;
Best Local Similarity	56.48;	Pred. No. 8.1e-24;		
Matches 190; Conservative	0;	Mismatches 147;	Indels 0;	Gaps 0;

QY	31	GCAGCAATTTGTGTGTGAAGAACATCAAGCGCCAGAAATTTGGACGCCGGAGATTGACA	90
Db	57	GCCCGAGATTACAAGTCAAGGACATGTCTCAGGCCGACTTCGGCCGCTCGAATATCAGC	118
QY	91	TTGCAGAGCAAGACATGTCTGTCTGTATTCACTCAGAGAAACGTGCTCAGGGGGAGAGC	150
Db	117	TGCGGTGAGGTGAGATGGCAGAGGGTCAATGGGCTCGCGTGCAGAAATTCGGCCCCCAGC	176
QY	151	CCTTGCGTGTGTCTAAATAATAGGGCTGTACACACATCACAGCCACAGACGGGTGTGA	210
Db	177	CATTCAGGGCGGCAAAAATCACTGGATCCCTCCACATGAGAGATCCAACACTCGCTCTCA	236
QY	211	TTGAGACACTGTGCCCCCTGGGGGCTCAGTACCCTGTGTGCTTTAAATCTACTCAA	270
Db	237	TGCAAAACCTTCAACGCGCTCGGGGCCAGAGTTCCCTGTGTCTCTGCAACATATTCCCA	296
QY	271	CTCAGAATGAAGTAGCTGCACACACTGGGTGAGGGCTGGAGTTGACAGTTTCGCTTGGAAAG	330
Db	297	CCCGGAGCAATGCGCGCGCTCCATTCGCCGTGACTCCGCTCCGTCTTGCGCTGGAGG	356
QY	331	GCGAGTCAAGAGATGACTTGTGTGTGTATTGACCG	367
Db	357	GTGAGACCTTCACAGAGTACTGTGTGTGTACACCGC	393

RESULT 6
US-08-204-740-8
Secured 8 1001484100 08/08204740

GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allgretti & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:

```

MEDTUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,740
FILING DATE: 04-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5753432nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-204-740-8

Query Match      17.4%; Score 72.6; DB 1; Length 289;
Best Local Similarity 58.6%; Pred.No.1,4e-14;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

OY 52 ACATTAACGACGCAAAATTGGAGCCCGGAGAAATTGAGATTGCAGAGCAAGCATGTCTG 111
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 75 ACATGGGCGCTGGCTCCTCGGGAGACGCCAAGGCCCTGACACTTGCTGAAGAAGATGCGG 134
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 112 CTCGATTTACTACACGAAAAGTCGCTCAGGGGAGAACGCCCTTGCTGCTAAATAAG 171
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 135 GCCTGATCGATGATGGGGGAGCGGTACTCGGCTCCAGGCCACAGTAGGGCGCCCGCATCG 194
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 172 TGGCGTGAACAACATCACAGCCGACAGACGGGGTGTGATTGAGACACTGTGCGCCTGG 231
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 195 CTGGCTGCGCTGACATGACCCTGGAGACGGCGCTCCTCATTTGAGACCCCTGTCACCCCTGG 254
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 232 GGGCTCAGTGGCCGTGTCTGCTGTTGAACATCTAC 266
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 255 GTGCTGAGGTGTCAGTGTCTGACACTGCACATCTTC 289
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-09-081-167A-8
Sequence 8, Application US/09081167A
Patent No. 6083745
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive, 32nd Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,167A

```

```

FLILING DATE: 18-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 6083746nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-KK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-081-167A-8

Query Match 17.4%; Score 72.6; DB 3; Length 289;
Best Local Similarity 58.6%; Pred. No. 1.4e-14;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 52 ACATCAGACAGCAGCAATTTGGACGCCGGAGATGAGATTGCAGAGCAAGACATCTCTG 111
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 75 ACATGGCGCTGCTGCTGCTGGGAGCGCAAGGCCCTGTGACATCTGTGAGAACGACATCGCG 134
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 112 CTCGTGATTTCACTCAGGAACGCTGCTCAGGGGGGAGAACCCCTTGCTGGTGTCTAAATAG 171
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 135 GCGCTATGCGTATGGGGGGAGCGGTACTCGGCCCTCCAGGCCACACGAGGAGGGGCCCGCATCG 194
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 172 TGGGGCTGTACACATCAGACGCCCGCAGACAGCGGTGTGATTGAGACATCTGTGCGCTGG 231
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 195 CTGGCTGCTGCACATGACCGCTGGAGACGCCGCTCTCATTTGAGACCCCTGTCACCCCTGG 254
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 233 GGGCTCAGTGCCTGCTGCTGCTGCTGTGATGACATCTAC 266
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 255 GTGCTGAGGTGCAGTGGTCCAGCTGCACATCTTC 289
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-09-081-395-8
: Sequence 8, Application US/09081395
: Patent No. 6083746
: GENERAL INFORMATION:
: APPLICANT: Gudkov, Andrei
: APPLICANT: Kazarov, Alexander
: APPLICANT: Mazo, Ilya
: APPLICANT: Roninson, Igor B
: TITLE OF INVENTION: Methods for Identifying Genetic
: TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
: TITLE OF INVENTION: Growth in Cancer Cells
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
: STREET: 300 S. Wacker Drive, 32nd Floor
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/081,395
: FILING DATE: 18-MAY-1998
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 6083746nan, Kevin E
: REGISTRATION NUMBER: 35,303
: REFERENCE/DOCKET NUMBER: 93,354-KK

```

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001

TELEFAX: 312-913-0002

TELEX:

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 289 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-09-081-395-8

Query Match

17.4%; Score 72.6; DB 3; Length 289;

Best Local Similarity 58.6%; Pred. No. 1.4e-14;

Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 52 ACATCAGCAGCAGCAATTGGACGCCGGAGATTGAGTTGCAGAGCAACATCTCTG 111
 DB 75 ACATCGCCCTGGCTGCTCGGGAGCGCAAGCCCTGGACATTTGCTGAGAACGAGATCCGG 134
 QY 112 CTCGTATTTCATCTCAGGAAACGTGCTCAGGGGGAGAGCCCTTGGTGGTCTAAATAG 171
 DB 135 GCCTATGCGTATGCGGGAGCGGTACTGCGCTCCAGCCACTGAAGGCGCCCGCATCG 194
 QY 172 TGGCGTGTACACATCAGACAGCCGAGAGCGGTGTGATTGAGACACTGTGCCCCCTG 231
 DB 195 CTGGCTGCTGCACATGACCGGTGAGAGCGCGCTCTCATTTAGAGACCTGTCACCCCTG 254
 QY 232 GGGCTCAGTGCCTGCTGCTGCTGCTGTTGATTAACATCTAC 266
 DB 255 GTGCTGAGGTGCAGTGTGCTCAGCTGCAACATCTTC 289

RESULT 9

US-09-416-833-8

Sequence 8, Application US/09416833

Patent No. 6197521

GENERAL INFORMATION:

APPLICANT: Gudkov, Andrei

APPLICANT: Kazarov, Alexander

APPLICANT: Mazo, Ilya

APPLICANT: Koninson, Igor B

TITLE OF INVENTION: Methods for Identifying Genetic

TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allgretti & Witcoff, Ltd.

STREET: 10 S. Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/416,833

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/204,740

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: No. 6197521man, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 93,354-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 289 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-09-416-833-8

Query Match

17.4%; Score 72.6; DB 4; Length 289;

Best Local Similarity 58.6%; Pred. No. 1.4e-14;

Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 52 ACATCAGCAGCAGCAATTGGACGCCGGAGATTGAGTTGCAGAGCAACATCTCTG 111
 DB 75 ACATCGCCCTGGCTGCTCGGGAGCGCAAGCCCTGGACATTTGCTGAGAACGAGATCCGG 134
 QY 112 CTCGTATTTCATCTCAGGAAACGTGCTCAGGGGGAGAGCCCTTGGTGGTCTAAATAG 171
 DB 135 GCCTATGCGTATGCGGGAGCGGTACTGCGCTCCAGCCACTGAAGGCGCCCGCATCG 194
 QY 172 TGGCGTGTACACATCAGACAGCCGAGAGCGGTGTGATTGAGACACTGTGCCCCCTG 231
 DB 195 CTGGCTGCTGCACATGACCGGTGAGAGCGCGCTCTCATTTAGAGACCTGTCACCCCTG 254
 QY 232 GGGCTCAGTGCCTGCTGCTGCTGCTGTTGATTAACATCTAC 266
 DB 255 GTGCTGAGGTGCAGTGTGCTCAGCTGCAACATCTTC 289

RESULT 10

PCT-US95-02521-8

Sequence 8, Application PC/TUS9502521

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Methods for Identifying Genetic

TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

NUMBER OF SEQUENCES: 13

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/02521

FILING DATE:

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 289 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

PCT-US95-02521-8

Query Match

17.4%; Score 72.6; DB 5; Length 289;

Best Local Similarity 58.6%; Pred. No. 1.4e-14;

Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 52 ACATCAGCAGCAGCAATTGGACGCCGGAGATTGAGTTGCAGAGCAACATCTCTG 111
 DB 75 ACATCGCCCTGGCTGCTCGGGAGCGCAAGCCCTGGACATTTGCTGAGAACGAGATCCGG 134
 QY 112 CTCGTATTTCATCTCAGGAAACGTGCTCAGGGGGAGAGCCCTTGGTGGTCTAAATAG 171
 DB 135 GCCTATGCGTATGCGGGAGCGGTACTGCGCTCCAGCCACTGAAGGCGCCCGCATCG 194
 QY 172 TGGCGTGTACACATCAGACAGCCGAGAGCGGTGTGATTGAGACACTGTGCCCCCTG 231
 DB 195 CTGGCTGCTGCACATGACCGGTGAGAGCGCGCTCTCATTTAGAGACCTGTCACCCCTG 254

QY 232 GGGCTCAGTCCGCGTGTCTGTTGTAACATCTAC 266
DB 255 GTGCTGAGGTGCGAGTGTCCAGCTGCACATCTTC 289

RESULT 11

US-08-204-740-6
Sequence 6, Application US/08204740
Patent No. 5753432
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allgretti & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,740
FILING DATE: 04-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5753432nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-204-740-6

Query Match 16.4%; Score 68.4; DB 1; Length 285;
Best Local Similarity 56.8%; Pred. No. 3.4e-13;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 45 GTGAGAACATCAAGCAGGAGAAATTTGAGCCGCGGAGATTGATGACAGCAAGAC 104
DB 64 GTCCGGACATCGGACGCGCTGCGGAGAGAGCTGTGATATAGCTGAGAAATGAG 123
QY 105 ATGCTCTCTGATTTCACACAGAAAGCTGACGAGGAGAGCCCTGGCGGTGCT 164
DB 124 ATGCGAGGTTGATGCGCATGCGGAGATGATCTACGCTCCAGCCACTGAGAGGTGCT 183
QY 165 AAATATGAGGCTGTACACATCAACATCAGACCCAGACAGCGGTGTTGATGAGACACTCTGT 224
DB 184 CGCATTTGCTGCTCCGCGCATACCGTGAAGACTGTGTTTCATATGAGACTCTGCTG 243
QY 225 GCCCTGGGGCTCAGTGCCGCTGCTGCTGTGTAACATCTAC 266
DB 244 GCCCTGGGTGCTGAGGCGGCTGTGTCAGCTGCAACATCTTC 285

RESULT 12

US-09-081-167A-6
Sequence 6, Application US/09081167A
Patent No. 6083745
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive, 32nd Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,167A
FILING DATE: 18-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 6083745nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-KK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-081-167A-6

Query Match 16.4%; Score 68.4; DB 3; Length 285;
Best Local Similarity 56.8%; Pred. No. 3.4e-13;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 45 GTGAGAACATCAAGCAGGAGAAATTTGAGCCGCGGAGATTGATGACAGCAAGAC 104
DB 64 GTCCGGACATCGGACGCGCTGCGGAGAGAGCTGTGATATAGCTGAGAAATGAG 123
QY 105 ATGCTCTCTGATTTCACACAGAAAGCTGACGAGGAGAGCCCTGGCGGTGCT 164
DB 124 ATGCGAGGTTGATGCGCATGCGGAGATGATCTACGCTCCAGCCACTGAGAGGTGCT 183
QY 165 AAATATGAGGCTGTACACATCAACATCAGACCCAGACAGCGGTGTTGATGAGACACTCTGT 224
DB 184 CGCATTTGCTGCTCCGCGCATACCGTGAAGACTGTGTTTCATATGAGACTCTGCTG 243
QY 225 GCCCTGGGGCTCAGTGCCGCTGCTGCTGTGTAACATCTAC 266
DB 244 GCCCTGGGTGCTGAGGCGGCTGTGTCAGCTGCAACATCTTC 285

RESULT 13

US-09-081-395-6
Sequence 6, Application US/09081395
Patent No. 6083746
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander


```

APPLICANT: Mazo, Ilya
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSER: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive, 32nd Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,395
FILING DATE: 18-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 6083746nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-KK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-081-395-6

Query Match 16.4%; Score 68.4; DB 3; Length 285;
Best Local Similarity 56.8%; Pred. No. 3,4e-13;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 45 GTGAAGACATCAAGCAGCAGATTTGACGCCGGGAGATTGAGATTGACAGCAAGAC 104
DB 64 GTGGCGGACATGAGCTGGCCCGCTGGGGAGCGAAGCTCTGATTTAGCTGAGATGAG 123
QY 105 ATGCTGCTGTGATTTCACTCAAGAAAGCTGCTCAAGGGGAGAGCCCTGGCTGCTGCT 164
DB 124 ATGCCAGGCTTATGCGCATGCGGGAGATGTACTGACCTCCAGGCACTGAAGGGTGTCT 183
QY 165 AAAATAGTGGGCTGTACACATCAACAGCCAGAGCCGCTGTGTTGATGAGACATCTGT 224
DB 184 CGCATTTGCTGGTGGCTGGCTGGCATGACCGGTGAGAGACTGCTGTCTCATTTAGACTCTCG 243
QY 225 GCCCTGGGGGCTCAGTCCCGCTGGTGTCTGTTGATACATCTTAC 266
DB 244 GCCCTGGGTGCTGAGGCGCGGTGTCTCAAGTGTCAACATCTTTC 285

RESULT 14
US-09-416-833-6
Sequence 6, Application US/09416833
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13

```

```

CORRESPONDENCE ADDRESS:
ADDRESSER: Allgretti & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/416,833
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/204,740
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 6197521nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-416-833-6

Query Match 16.4%; Score 68.4; DB 4; Length 285;
Best Local Similarity 56.8%; Pred. No. 3,4e-13;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 45 GTGAAGACATCAAGCAGCAGATTTGACGCCGGGAGATTGAGATTGACAGCAAGAC 104
DB 64 GTGGCGGACATGAGCTGGCCCGCTGGGGAGCGAAGCTCTGATTTAGCTGAGATGAG 123
QY 105 ATGCTGCTGTGATTTCACTCAAGAAAGCTGCTCAAGGGGAGAGCCCTGGCTGCTGCT 164
DB 124 ATGCCAGGCTTATGCGCATGCGGGAGATGTACTGACCTCCAGGCACTGAAGGGTGTCT 183
QY 165 AAAATAGTGGGCTGTACACATCAACAGCCAGAGCCGCTGTGTTGATGAGACATCTGT 224
DB 184 CGCATTTGCTGGTGGCTGGCTGGCATGACCGGTGAGAGACTGCTGTCTCATTTAGACTCTCG 243
QY 225 GCCCTGGGGGCTCAGTCCCGCTGGTGTCTGTTGATACATCTTAC 266
DB 244 GCCCTGGGTGCTGAGGCGCGGTGTCTCAAGTGTCAACATCTTTC 285

RESULT 15
PCR-US95-02521-6
Sequence 6, Application PC/TUS9502521
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02521

```



```
Db 649 CACTCAGGAACGCTGCTCAGAGGAGAGCCCTGCTGCTGCTAAATAGTGGCTGTA 708
181 CACACATCACAGCCACAGACAGCGCTGTTGATTGAGACACTCTGCCCCCTGGGGCTCAGT 240
709 CACACATCACAGCCACAGACAGCGCTGTTGATTGAGACACTCTGCCCCCTGGGGCTCAGT 768
241 GCCGCTGCTGCTGTTGATTGAGACACTCTGAGAAATGAGTAGTGCAGCAGCTGCTG 300
769 GCCGCTGCTGCTGTTGATTGAGACACTCTGAGAAATGAGTAGTGCAGCAGCTGCTG 828
301 AGCGTGAGTGTGAGTGTGCTGTTGAGAGGCGAGTGCAGAAAGATGCTGCTGCTGTA 360
829 AGCGTGAGTGTGAGTGTGCTGTTGAGAGGCGAGTGCAGAAAGATGCTGCTGCTGTA 888
361 TTGACCGCTGTGTGACATGATGATGGTGTGCGAGGCGCAACATGATCCTGATGATGGG 417
889 TTGACCGCTGTGTGACATGATGATGGTGTGCGAGGCGCAACATGATCCTGATGATGGG 945
```

RESULT 2

```
US-09-879-536-850/c
; Sequence 850, Application US/09879536
; Patent No. US20020144298A1
```

GENERAL INFORMATION:

```
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll, III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Dertl, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/879, 536
CURRENT FILING DATE: 2001-09-21
PRIORITY FILING DATE: 1998-06-10
PRIORITY FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 850
LENGTH: 636
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(636)
OTHER INFORMATION: n - A,T,C or G
US-09-879-536-850
```

```
Query Match 54.3%; Score 226.4; DB 10; Length 636;
Best Local Similarity 99.2%; Pred. No. 3.6e-67;
Matches 238; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
```

```
Db 179 TACACATACACAGCCACAGACAGC-GGTGTTGATTGAGACACTCTGCGCCCTGGGGCTC 237
420 TACACATACACAGCCACAGACAGC-GGTGTTGATTGAGACACTCTGCGCCCTGGGGCTC 361
238 AGTCCCGCTGCTGCTGTTGATTGAGACACTCTGAGAAATGAGTAGTGCAGCAGCTG 297
360 AGTCCCGCTGCTGCTGTTGATTGAGACACTCTGAGAAATGAGTAGTGCAGCAGCTG 301
298 CTGAGGCTGAGTGTGAGTGTGCTGTTGAGAGGCGAGTGCAGAAAGATGATCTTGGTGT 357
300 CTGAGGCTGAGTGTGAGTGTGCTGTTGAGAGGCGAGTGCAGAAAGATGATCTTGGTGT 241
358 GTATTGACCGCTGTGTGACATGATGATGGTGTGCGAGGCGCAACATGATCTTGGATGATGGG 417
```

```
Db 240 GTATTGACCGCTGTGTGACATGATGATGGTGTGCGAGGCGCAACATGATCTTGGATGATGGG 181
```

RESULT 3

```
US-09-925-300-439
; Sequence 439, Application US/09925300
; Patent No. US20020151681A1
```

GENERAL INFORMATION:

```
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIORITY FILING DATE: 2000-03-08
PRIORITY FILING DATE: 2000-03-08
PRIORITY FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 439
LENGTH: 721
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (688)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-439
```

```
Query Match 42.7%; Score 178; DB 10; Length 721;
Best Local Similarity 94.2%; Pred. No. 1.3e-50;
Matches 178; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
```

```
Db 1 GAGAGAAGCAGCAACCACTCAAGGCGAGCAGCAATTTCTGTGTGAAGACATCAAGC 60
532 GAGAGAAGCAGCAACCACTCAAGGCGAGCAGCAATTTCTGTGTGAAGACATCAAGC 591
61 AGGCAATTTGAGCGCGGAGATTGAGATTGAGAGCAAGACATGCTCTCTGATTT 120
592 AGGCAATTTGAGCGCGGAGATTGAGATTGAGAGCAAGACATGCTCTCTGATTT 651
121 CACTCAGGAACGCTGCTCAGGCGGAGAGACCCCTGCGGTGCTAAATAGTGGGCTGTA 180
652 CACTCAGGAACGCTGCTCAGGCGGAGAGACCCCTGCGGTGCTAAATAGTGGGCTGTA 711
181 CACACATCA 189
712 CACATACA 720
```

RESULT 4

```
US-09-925-301-217
; Sequence 217, Application US/09925301
; Patent No. US20020052308A1
```

GENERAL INFORMATION:

```
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIORITY FILING DATE: 2000-03-08
PRIORITY FILING DATE: 2000-03-08
PRIORITY FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 217
LENGTH: 2200
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
```

NAME/KEY: misc_feature
LOCATION: (2188)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-217

Query Match 32.9%; Score 137.2; DB 10; Length 2200;
Best Local Similarity 60.9%; Pred. No. 1.7e-36;
Matches 223; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 52 ACATCAAGCAGGAGATTTGACGCCGGAGATGTGATGACAGACACATGCTCTG 111
DB 108 ACATGCGCCCTGGCTGCTGGGAGCAGAGCCCTGACATGCTGAGAACAGAGATGCGG 167
QY 112 CTGTGATTTCACTCAGGAAACGTGCTCAGGGGAGAGAGCCCTTGCTGGCTGCTAAATAG 171
DB 168 GCTGTATGCGTATGCGGAGCGGTACTCGGCTCCAGACCTGAGAGGCGCCGCTCG 227
QY 172 TGGGCTGTACACATCACACCCAGACAGCGGTGTGATGAGACATCTGTGCTCTG 231
DB 228 CTGCTGCGCTGACATGACCGGTGAGAGCGCCGCTCATTTGAGACCCCTGTCACCTGG 287
QY 232 GGGCTAGTGGCCCTGCTGCTGTGATGATCTACTCACTCAATGAGATGAGTCTGAG 291
DB 288 GTGCTAGGTGAGTGTGCTGCTGAGCAACATCTCTCCACCCAGGACATGCGCGCTG 347
QY 292 CACTGCTGAGCGTGGAGTGTGAGTGTGCTGAGAGGCGAGTCAAGATGACTTCT 351
DB 348 CCAATGCGCAAGGCTGCGATTCGCTGTATGCTGAGAGGCGAGACGAGAGATGACC 407
QY 352 GTGTGTATTTGACCGCTGTGTGATGATGATGATGATGATGATGATGATGATGATG 411
DB 408 TGTGTGTATTTGACAGACCGCTGTACTTCAAGAGAGGCGCCCTCAACATGATCTGAGC 467
QY 412 ATGGGG 417
DB 468 ACGGGG 473

RESULT 5
US-10-044-090-344
Sequence 344, Application US/10044090
Patent No. US20020137081A1

GENERAL INFORMATION:

APPLICANT: Olga Bandman
TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
FILE REFERENCE: PA-0028 US
CURRENT APPLICATION NUMBER: US/10/044, 090
CURRENT FILING DATE: 2002-01-09
NUMBER OF SEQ ID NOS: 850
SOFTWARE: PERL Program
SEQ ID NO 344
LENGTH: 2429
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020137081A1 1468237CBI
US-10-044-090-344

Query Match 32.9%; Score 137.2; DB 12; Length 2429;
Best Local Similarity 60.9%; Pred. No. 1.8e-36;
Matches 223; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 52 ACATCAAGCAGGAGATTTGACGCCGGAGATGTGATGACAGACACATGCTCTG 111
DB 108 ACATGCGCCCTGGCTGCTGGGAGCAGAGCCCTGACATGCTGAGAACAGAGATGCGG 167
QY 112 CTGTGATTTCACTCAGGAAACGTGCTCAGGGGAGAGAGCCCTTGCTGGCTGCTAAATAG 171
DB 168 GCTGTATGCGTATGCGGAGCGGTACTCGGCTCCAGACCTGAGAGGCGCCGCTCG 227
QY 172 TGGGCTGTACACATCACACCCAGACAGCGGTGTGATGAGACATCTGTGCTCTG 231
DB 468 ACGGGG 473

DB 228 CTGCTGCTGCAACATGACCGGTGAGAGCGCCGCTCTCATTTAGACCCCTGTCACCTGG 287
QY 232 GGGCTAGTGGCCCTGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATG 291
DB 288 GTGCTAGGTGAGTGTGCTGCTGCAACATCTCTCCACCCAGGACATGCGCGCTG 347
QY 292 CACTGCTGAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 351
DB 348 CCAATGCGCAAGGCTGCGATTCGCTGTATGCTGAGAGGCGAGAACGAGAGATGACC 407
QY 352 GTGTGTATTTGACCGCTGTGTGATGATGATGATGATGATGATGATGATGATGATG 411
DB 408 TGTGTGTATTTGACAGACCGCTGTACTTCAAGAGAGGCGCCCTCAACATGATCTGAGC 467
QY 412 ATGGGG 417
DB 468 ACGGGG 473

RESULT 6
US-09-966-881-9
Sequence 9, Application US/09966881
Patent No. US20020120960A1

GENERAL INFORMATION:

APPLICANT: Seymour, Graham
Blind, Collin
Medina-Suarez, Rosybel
TITLE OF INVENTION: Genetic Control of Fruit Ripening
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zeneca Ag Products Inc.
STREET: 1800 Concord Pike
CITY: Wilmington
STATE: DE
COUNTRY: USA
ZIP: 19850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/966, 881
FILING DATE: 28-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/242, 860
FILING DATE: 29-Mar-1999
APPLICATION NUMBER: GB 9618862.8
FILING DATE: 10-Sep-1996
APPLICATION NUMBER: GB 9708366.1
FILING DATE: 25-Apr-1997
APPLICATION NUMBER: PCT/GB97/02424
FILING DATE: 08-Sep-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hohnsenschutz, Liza D.
REGISTRATION NUMBER: 33,712
REFERENCE/DOCKET NUMBER: SEE 50183/UST
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 886-1699
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 793 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: U-0131
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-966-881-9

Query Match 27.4%; Score 114.4; DB 10; Length 793;

Best Local Similarity 58.0%; Pred. No. 6.6e-29;
Matches 202; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 34 GCAATTTCTGTGTAAGACATCAAGCAGCATTTGGACGCCGGAGATTAGATTG 93
DB 82 GCGAGTACAAAGGTGAGACCTTTCTCAGGCTGACTTCGGCGCTCGAGTCAAGCTGG 141
QY 94 CAGAGCAAGACATCTGCTGCTGATTTCACTCAGAAACGTGCTCAGGGGAGAGACCTT 153
DB 142 CCGAGGTGAGATGCGGGCTCTCATGGCTGCCGCCGAGATTGGGGCCCAAACTT 201
QY 154 TGCGTGTGCTAAATAGTGGCTGTACACATCACAGCCAGACAGCGGTGTGATTG 213
DB 202 TGCGGGGGGGCGGATCTCCGGCTCCCTCCACATGACATCCAGACCGCTCCATCG 261
QY 214 AGACACTGTGCTTGGGGGCTCAGTCCGCTGCTGTGTAACATCTACTCACTC 273
DB 262 AGACCTTCACCGCCCTGCTGCTCAAGTCCGCTGCTGTGTAACATCTACTCACTC 321
QY 274 AGAATGAGTGTGCTGCAAGCACTGGCTGAGCTGAGTGTGCAATGCTTGGAAAGGCG 333
DB 322 AGACACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 381
QY 334 AGTCAAGATGACTTCTGTGTGTATTGACCGCTGTGTGTAACATG 381
DB 382 AGACCTTCGCGGAGTACTGCTGTGCTGCAACCGAGCATGCTGACTGGG 429

RESULT 7

US-10-037-598-26
Sequence 26, Application US/10037598
Patent No. US20020157143A1
GENERAL INFORMATION:
APPLICANT: Monsanto Co
APPLICANT: Concilio, Verge
APPLICANT: Delaney, Xavier
TITLE OF INVENTION: Soybean Plants with Enhanced Yields and Methods for Breeding for
TITLE OF INVENTION: Screening of Soybean Plants with Enhanced Yields
FILE REFERENCE: 38-21(52175)B
CURRENT APPLICATION NUMBER: US/10/037, 598
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 06/260,040
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patent version 3.0
SEQ ID NO 26
LENGTH: 3830
TYPE: DNA
ORGANISM: Glycine max
US-10-037-598-26

Query Match 27.4%; Score 114.4; DB 9; Length 3830;
Best Local Similarity 57.1%; Pred. No. 1.3e-28;
Matches 208; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 18 AACTCCAGGCGAGCAATTTCTGTGTAAGACATCAAGCAGCATTTGGAGC 77
DB 1753 AAAACACAGAGTGTGCGAGATCAAGGTCAAGGACCTTCCAGGCGGACTTGGCGC 1812
QY 78 CCGGAGATTGATGAGACATCAAGACATGCTGCTGATTTCACTCAGTCAAGAGTGTCT 137
DB 1813 CTCGAGATCGAGCTGCGGAGGTGAGATGCCCGGCTCAATGCTGTGCGAGCATTC 1872
QY 138 CAGGGGAGAGCCCTTGGCTGTCTAAATAGTGGCTGTACACATCAAGCAGCCAG 197
DB 1873 GCGCCCTCCAGCCCTTCAAGGGGCGCGCATCACCGGCTCCCTCCATGACATTCAG 1932
QY 188 ACAGGAGTGTGATGAGACATCTGTGCTGCGGGGCTCAATGCGCGCTGCTGTCTGT 257
DB 1933 ACCGCGCTTCTCATTTGAGACCTTCACCGCCTTGGCGGAGTCCGCTGTGCTCTGC 1992
QY 258 AACATCTACTCAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 317

DB 1993 AACATCTCTCCACCAGAGACCGCCGCCCTATTGCCCGGACAGTGGCGCTC 2052
QY 318 TTGCTTGGAGGGCGAGTCAAGATGATCTTGTGTGTGATGACCGCTGTGTGAC 377
DB 2053 TTGCTTGGAGGGTGAAGACCTCCAGAGTACTGTGTGTGTGTGTGTGTGTGTGTGT 2112
QY 378 ATGG 381
DB 2113 TGGG 2116

RESULT 8

US-09-754-853A-4/c
Sequence 4, Application US/09754853A
Publication No. US20030005491A1
GENERAL INFORMATION:
APPLICANT: Hauge, Brian M.
APPLICANT: Parnell, Laurence D.
APPLICANT: Parsons, Jeremy D.
APPLICANT: Mang, Ming Li
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Soybean Cyst Nematode Resistance
FILE REFERENCE: 38-10(15810)B
CURRENT APPLICATION NUMBER: US/09/754, 853A
PRIOR FILING DATE: 2001-01-05
PRIOR APPLICATION NUMBER: US 60/174,880
NUMBER OF SEQ ID NOS: 1119
SEQ ID NO 4
LENGTH: 513509
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: CDS
LOCATION: (11805)..(113968)..(114684)..(115204)
NAME/KEY: unsure
LOCATION: (1)..(513509)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: clone id: 318013_region_A3
US-09-754-853A-4

Query Match 27.4%; Score 114.4; DB 9; Length 513509;
Best Local Similarity 57.1%; Pred. No. 1e-27;
Matches 208; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 18 AACTCCAGGCGAGCAATTTCTGTGTAAGACATCAAGCAGCATTTGGAGC 77
DB 180287 AAAACACAGAGTGTGCGAGATCAAGGTCAAGGACCTTCCAGGCGGACTTGGCGC 180228
QY 78 CCGGAGATTGATGAGACATCAAGACATGCTGCTGATTTCACTCAGTCAAGAGTGTCT 137
DB 180227 CTCGAGATCGAGCTGCGGAGGTGAGATGCCCGGCTCATGCTGTGCGAGCATTC 180168
QY 138 CAGGGGAGAGCCCTTGGCTGTCTAAATAGTGGCTGTACACATCAAGCAGCCAG 197
DB 180167 GCGCCCTCCAGCCCTTCAAGGGGCGCGCATCACCGGCTCCCTCCATGACATTCAG 180108
QY 198 ACAGGAGTGTGATGAGACATCTGTGCTGCGGGGCTCAATGCGCGCTGCTGTCTGT 257
DB 180107 ACCGCGCTTCTCATTTGAGACCTTCACCGCCTTGGCGGAGTCCGCTGTGCTCTGC 180048
QY 258 AACATCTACTCAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 317
DB 180047 AACATCTCTCCACCAGAGACCGCCCGGCTGATTTGCGGAGCATGCTCCCGC 179988
QY 318 TTGCTTGGAGGGCGAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATG 377
DB 179987 TTGCTTGGAGGGTGAAGACCTCCAGAGTACTGTGTGTGTGTGTGTGTGTGTGTGT 179928
QY 378 ATGG 381
DB 179927 TGGG 179924


```

US-09-746-660A-97
; Sequence 97, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krieger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberer, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/746,660A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 97
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1534)
; OTHER INFORMATION: RXN00132
US-09-746-660A-97

Query Match          27.1%; Score 113; DB 9; Length 1557;
Best Local Similarity 59.4%; Pred. No. 2.6e-28;
Matches 223; Conservative 0; Mismatches 130; Indels 24; Gaps 1;

63 GCAGATTGGACGCGGAGATTGAGATTGACAGCAAGCATGTGCTGATTTCA 122
143 GCAGAGGACGACGCTACACGATTGCTTGACAGATGAGATGCCAGGCTCATG 202
123 CTCAGAAAGCTGCTCAGGGGAGAACCCCTTGCTGCTAAATAGTGGCTGTACA 182
203 TTGGCGAAGGAATTCGAGAGACAGCCTTTGAAAGGCGCCGAATTCGTTCTATC 262
183 CACATCAGACCCAGACAGGGGTGTTGATTGACACTCTGTGCTGGGGGCTCACTG 242
263 CACATGAGGCTGACGAGCCGCTGCTTATGAGACCTCACTGTTGGGGGCTGAGGTT 322
243 CGGTGGCTGCTTATCACTACTCACTCAAGTGAAGTGAAGTGAAGGAGGCTGCTG 300
323 CGTGGGCTTCTGCAACATTTTCCACCCAGATGAGGTGAGGGGCTATGTTGTC 382
301 -----AGCTGAGATTGACGTTGCTGCTTGGAGGGGAGTCA 338
383 GCGTCGCGACCGTCGAAAGAGCCAGCTGGTTCACATATTCGGGTGGAAGGGTGA 442
339 GAGATACCTTCTGCTGATTTGACCGCTGTGTAACATGATGGGTGGCAGGCCAAC 398
443 CTGAGAGAGTACTGCTGCTGATCACCAGATCTTCAGCTGGGGCGATGAGCTCCAAAC 502
399 ATGATCTGATGATGGGG 417

```

```

Db          503 ATGATCTGACGACGCGG 521

RESULT 12
US-09-960-352-10073
; Sequence 10073, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nageshan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 10073
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (284)..(292)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 43-LIB34-036-01-El-C4
US-09-960-352-10073

Query Match          25.9%; Score 108.2; DB 10; Length 389;
Best Local Similarity 61.1%; Pred. No. 6.4e-27;
Matches 173; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

52 ACATCAAGCAGCAAAATTGGACCGCGGAGATTGAGATTGACAGCAGACATGCTG 111
105 ACATCAGCTGGCGCTGGGAGCAGAGCCCTGACCTGCAAGAAATGATGATGCGG 164
112 CTCGATTCTACCTAGGAACCTGCTCAGGGGAGAACCCCTGCTGCTAAATAG 171
165 GCGTATGACATCGGAATGATGCTGCGCTCCAGCCCTGAAAGGCGCTGCATG 224
172 TGGGCTGTACACATCAGACAGCCAGAGCGGTGATTGAGACACTGTGCTGCTG 231
225 CTGCTGCTGCTGACATGACCTGAGACCGCGCTCATTTGAAACCTGCTGCTG 284
232 GGGCTGAGTGGCTGCTGCTGCTGTTAATCACTCACTCAGATGAGATGATGCTG 291
285 GTGCTGANGTGGGTGCTGCTGCTGCAATATCTCTCCACCCAGACATCAGACCTG 344
292 CACTGCTGAGGCTGAGATTGCAATGTTGCTTGGAGGGCGA 334
345 CCAATTGCCAAGGCTGGCATTCAGTGTACGCTGGAAGGGTGA 387

RESULT 13
US-09-738-626-837/C
; Sequence 837, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIRO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18

```


PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 837
LENGTH: 708
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-837

Query Match 25.6%; Score 106.8; DB 9; Length 708;
Best Local Similarity 59.4%; Pred. No. 2.5e-26;
Matches 214; Conservative 0; Mismatches 122; Indels 24; Gaps 1;

QY 82 AGATTGAGTTGACAGACAGACATGCTGCTGATTCACACAGGAAACGTCAGG 141
DB 706 AGATTGCTCTTCAGAGATGAGATGCGAGGCTCATGAGTTGCCAAGGAATTCGCA 647
QY 142 GGGAGAGCCCTTGCTGCTGCTAAATAGTGGCTGTACACACATCAGCCAGACAG 201
DB 646 AGGACAGCCTTTGAAAGGCGCCCAATGCTGCTCTATCCACATGAGGCTCCAGCCG 587
QY 202 CGGTGTTGATGAGACACTGTGTGCTGCGGAGCTCAGTCCGCTGCTGTGTTGTA 261
DB 586 CCGTCTTATTTAGACCCCTCACTGCTTTGGGCGCTGAGGTGCTGGCTTCTCGAACA 527
QY 262 TCTACTCAACATGAGATGAGTGTGACGACACTGCTG----- 300
DB 526 TTTTTCACCCAGATGAGTGGCTGACGCGCTATGCTTTCGCGCTCCGACCGCTGGAAG 467
QY 301 --AGCTGAGATTGACAGTGTGCTTGTGAGAGGCGAGACAGAGATGACTTGTGTG 357
DB 466 AGCCAGCTGCTTTCAGATTCGCGCTGGAAGGAGTCACTGAGAGGACTGTGCTG 407
QY 358 GTATTGACCGCTGTGTGAACATGATGAGGTGCGAGCCACATGATCTGATGATGGG 417
DB 406 GCATCAACCAAGTCTTCAGCTGCGGCGATGAGCTGCACAAATGATCCCTGAGAGGGG 347

RESULT 14
US-09-770-149-117/c

Sequence 117, Application US/09770149

Patent No. US20020059663A1

GENERAL INFORMATION:

APPLICANT: Gorlach, Jörn

APPLICANT: An, Jong-Oladg

APPLICANT: Hamilton, Carol M.

APPLICANT: Price, Jennifer L.

APPLICANT: Raines, Tracy M.

APPLICANT: Yu, Yang

APPLICANT: Rameaka, Joshua G.

APPLICANT: Page, Amy

APPLICANT: Matthew, Abraham V.

APPLICANT: Ledford, Brooke L.

APPLICANT: Moessner, Jeffrey P.

APPLICANT: Haas, William David

APPLICANT: Garcia, Carlos A.

APPLICANT: Kriker, Maia

APPLICANT: Slader, Ted

APPLICANT: Davis, Keith R.

APPLICANT: Allen, Keith

APPLICANT: Hoffman, Neil

APPLICANT: Hurban, Patrick

TITLE OF INVENTION: Expressed sequences of Arabidopsis

FILE REFERENCE: 2024 (PARA-013PRV)

CURRENT APPLICATION NUMBER: US-09-770-149

CURRENT FILING DATE: 2001-01-26

PRIOR APPLICATION NUMBER: 60/178-506

PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 117
LENGTH: 728
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-149-117

Query Match 23.3%; Score 97.2; DB 10; Length 728;
Best Local Similarity 54.1%; Pred. No. 4.7e-23;
Matches 198; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 18 AACTCAAGGCGACAGCAATTTCTGTGTAAGACATCAAGCAGCAGATTGGACGC 77
DB 442 AAGACCTCAAGTGGCGGTGAATACAGGTCAAGACATGCTCAAGCCATTGCGCT 383
QY 78 CGGAGATTGAGATTGACAGCAAGACATGCTGCTGTGATTTCACTCAGAAACGCT 137
DB 382 CTGGAACCTGAGCTCCCGAAGTTGAGATGCCGTGACTCATGCTTGTCCGAAATTC 323
QY 138 CAGGGGAGAGCCCTTGCTGCTAAATAGTGGCGCTACACATCAGACGCCAG 197
DB 322 GACCTCTTCAGCATTCAGAGCGCTGATGATCAGGATCTTTCATGATGATCCAA 263
QY 198 ACAGCGTGTGTTGATGAGACACTGTGCTGCGGAGCTCAGTCCGCTGCTGCTGT 257
DB 262 ACCGCGTACTCATGAAACCCATGACTGCTGCTGCTGCTGATGATGATGCTTCTGC 203
QY 258 AACATCTACTCAACTGAGATGAGTACTGACGACACTGCTGAGGCTGAGTTGAGTG 317
DB 202 AACATCTTTCACCTCAAGCCAGCGCGCGCATGCTGCTGATCTCCGCGCTGTT 143
QY 318 TTTGCTGGAAGGCGAGTCAAGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 377
DB 142 TTGCGCTGGAAGTGAAGTACTTTCAGAGTACTGCTGCTGCTGCTGCTGCTGCTG 83
QY 378 ATGGAT 383
DB 82 TGGGGT 77

RESULT 15
US-09-960-352-4421

Sequence 4421, Application US/09960352

Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Ningbing

APPLICANT: Byatt, John C.

APPLICANT: Mathalagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AN

FILE REFERENCE: 16511.006/37-21(10298)C

CURRENT APPLICATION NUMBER: US/09/960,352

CURRENT FILING DATE: 2001-09-24

NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 4421

LENGTH: 433

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 19-LIB34-025-Q1-E1-E3

US-09-960-352-4421

Query Match 22.7%; Score 94.6; DB 10; Length 433;
Best Local Similarity 66.3%; Pred. No. 2.9e-22;
Matches 136; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 213 GAGACATCTGTGCGCTGAGGCTCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 272
DB 37 GAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 96
QY 273 CAGAATGAAGTAGTGCAGCACTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 332

Db 97 CAGGACCATGCAAGCAGCTGCCATTGCCAAGGCTGGCATTCAGTGTACGCTGGAGGGT 156
QY 333 GAGTCAGAGATGACTTCTGGTGGTATGACCGCTGTGACATGATGGGTGGCAG 392
Db 157 GAAAGGATGAGGATGACTGTGGTGCATGACAGACGCTGTACTTCAGGACGGGCC 216
QY 393 GCCAACATGATCCTGGATGATGGG 417
Db 217 CTCACATGATCTGAGACGCGTG 241

Search completed: April 21, 2003, 07:21:15
Job time : 336.47 secs

RESULT 1	
AU133884	
LOCUS	700 bp mRNA linear EST 01-AUG-2002
DEFINITION	AU133884 OVARC1 Homo sapiens cDNA clone OVARC100652 5', mRNA
ACCESSION	sequence.
VERSION	AU133884
KEYWORDS	AU133884.1 GI:10994423
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T. HRI human cDNA project Unpublished (2000) Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kusarazu, Chiba 292-0812, Japan Tel.: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomicehri.co.jp HRI human cDNA project; 5' - & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES
SOURCE

Location/Qualifiers
1. 700
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="OVARC100852"
/tissue_type="ovary, tumor tissue"
/note="Vector: pME18SPL3"

BASE COUNT 178 a 154 c 197 g 168 t 3 others
ORIGIN

Query Match 100.0%; Score 417; DB 9; Length 700;
Best Local Similarity 100.0%; Pred. No. 2,3e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAAACGACCAACCACTCCAGGCGACGACCAATTTCTGTGTGAAGACATCAAGC 60
DB 93 GAGGAAACGACCAACCACTCCAGGCGACGACCAATTTCTGTGTGAAGACATCAAGC 152
QY 61 AGGAGAAATTTGAGCGCGGAGATTGATTGAGAGCAAGACATGCTGCTGATTT 120
DB 153 AGGAGAAATTTGAGCGCGGAGATTGATTGAGAGCAAGACATGCTGCTGATTT 212
QY 121 CACTCAGGAAACGCTCTCAGGGGAGAAAGCCCTTGCTGCTAAATAGTGGCTGTA 180
DB 213 CACTCAGGAAACGCTCTCAGGGGAGAAAGCCCTTGCTGCTAAATAGTGGCTGTA 272
QY 181 CACACATACAGCCGACGAGGCTGATTGAGACACTGCTGCGGCGCTCAGT 240
DB 273 CACACATACAGCCGACGAGGCTGATTGAGACACTGCTGCGGCGCTCAGT 332
QY 241 GCCCGTGTCTGCTTGTAACTACTCACTCACTCAAGATGAGTCACTGAGCTGCTG 300
DB 333 GCCCGTGTCTGCTTGTAACTACTCACTCACTCAAGATGAGTCACTGAGCTGCTG 392
QY 301 AGGCTGAGTTCAGAGTGTCTGCTGGAAGGCGAGTCAAGAGATGATCTTGTGTGTA 360
DB 393 AGGCTGAGTTCAGAGTGTCTGCTGGAAGGCGAGTCAAGAGATGATCTTGTGTGTA 452
QY 361 TTGACCGCTGTGTAACTAGATGGTGGAGGCGCAACATGATCCTGATGATGGGG 417
DB 453 TTGACCGCTGTGTAACTAGATGGTGGAGGCGCAACATGATCCTGATGATGGGG 509

RESULT 2 731 bp mRNA linear EST 22-MAY-2001
BG826655
LOCUS 602748977F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4901851 5',
DEFINITION mRNA sequence.
ACCESSION BG826655
VERSION BG826655.1 GI:14174242
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 731)
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM1797 row: 1 column: 20
High quality sequence stop: 728.
Location/Qualifiers
1. 731

FEATURES
SOURCE

Location/Qualifiers
1. 700
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4901851"
/clone_id="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pORF7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 185 a 161 c 210 g 173 t 2 others
ORIGIN

Query Match 100.0%; Score 417; DB 12; Length 731;
Best Local Similarity 100.0%; Pred. No. 2,4e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAAACGACCAACCACTCCAGGCGACGACCAATTTCTGTGTGAAGACATCAAGC 60
DB 96 GAGGAAACGACCAACCACTCCAGGCGACGACCAATTTCTGTGTGAAGACATCAAGC 155
QY 61 AGGAGAAATTTGAGCGCGGAGATTGATTGAGAGCAAGACATGCTGCTGATTT 120
DB 156 AGGAGAAATTTGAGCGCGGAGATTGATTGAGAGCAAGACATGCTGCTGATTT 215
QY 121 CACTCAGGAAACGCTCTCAGGGGAGAAAGCCCTTGCTGCTAAATAGTGGCTGTA 180
DB 216 CACTCAGGAAACGCTCTCAGGGGAGAAAGCCCTTGCTGCTAAATAGTGGCTGTA 275
QY 181 CACACATACAGCCGACGAGGCTGATTGAGACACTGCTGCGGCGCTCAGT 240
DB 276 CACACATACAGCCGACGAGGCTGATTGAGACACTGCTGCGGCGCTCAGT 335
QY 241 GCCCGTGTCTGCTTGTAACTACTCACTCACTCAAGATGAGTCACTGAGCTGCTG 300
DB 336 GCCCGTGTCTGCTTGTAACTACTCACTCACTCAAGATGAGTCACTGAGCTGCTG 395
QY 301 AGGCTGAGTTCAGAGTGTCTGCTGGAAGGCGAGTCAAGAGATGATCTTGTGTGTA 360
DB 396 AGGCTGAGTTCAGAGTGTCTGCTGGAAGGCGAGTCAAGAGATGATCTTGTGTGTA 455
QY 361 TTGACCGCTGTGTAACTAGATGGTGGAGGCGCAACATGATCCTGATGATGGGG 417
DB 456 TTGACCGCTGTGTAACTAGATGGTGGAGGCGCAACATGATCCTGATGATGGGG 512

RESULT 3 744 bp mRNA linear EST 07-MAY-2001
BG699160
LOCUS 602678818F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4811477 5',
DEFINITION mRNA sequence.
ACCESSION BG699160
VERSION BG699160.1 GI:13967178
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 744)
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: Marios Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: LLAM10702 row: d column: 06
High quality sequence stop: 741.
Location/Qualifiers

FEATURES

source

1. 744
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:481477"
/clone_1lb="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: Brain; Vector: pBluescriptR (modified pBluescript KS+); Site: 1: BamHI; Site: 2: SalI-XhoI (gtcgag); Oligo-dt primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.5 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT

194 a 175 c 217 g 158 t

ORIGIN

Query Match 100.0%; Score 417; DB 12; Length 744;
Best Local Similarity 100.0%; Pred. No. 2.4e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGAGAGAGCAACCAATCTCCAGGCGAGCAATTTCTGTGTGAAGAATCAAGC 60
DB 241 GAGAGAGAGCAACCAATCTCCAGGCGAGCAATTTCTGTGTGAAGAATCAAGC 300
OY 61 AGGAGAAATTTGAGCGCGGAGATTGAGATTGAGAGCAAGACATGCTGCTGATT 120
DB 301 AGGAGAAATTTGAGCGCGGAGATTGAGATTGAGAGCAAGACATGCTGCTGATT 360
OY 121 CACTCAGGAAACGTGCTCAGGCGAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTA 180
DB 361 CACTCAGGAAACGTGCTCAGGCGAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTA 420
OY 181 CACACATCAGCCAGCCAGCGGCTGTGATTGAGACACTGCTGCGCTGGGCGCTAGT 240
DB 421 CACACATCAGCCAGCCAGCGGCTGTGATTGAGACACTGCTGCGCTGGGCGCTAGT 480
OY 241 GCGGCTGCTGCTGTTGAACATCTACTCACTCAGATGAGATGAGTCACTGCTG 300
DB 481 GCGGCTGCTGCTGTTGAACATCTACTCACTCAGATGAGATGAGTCACTGCTG 540
OY 301 AGGCTGAGATTGAGAGTGTGCTTGAAGGCGAGAGTCAAGATGACTTCTGTGTGTA 360
DB 541 AGGCTGAGATTGAGAGTGTGCTTGAAGGCGAGAGTCAAGATGACTTCTGTGTGTA 600
OY 361 TTGACCGCTGTGTGAACATGATGATGGTGAGGCGCAACATGATCTGTGATGATGGG 417
DB 601 TTGACCGCTGTGTGAACATGATGATGGTGAGGCGCAACATGATCTGTGATGATGGG 657

RESULT 4 763 bp mRNA linear EST 07-MAY-2001

LOCUS BG706498 60267003371 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792813 5', mRNA sequence.

ACCESSION BG706498
VERSION BG706498.1 GI:13981903

KEYWORDS

EST.

human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (Bases 1 to 763)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.fremail.nih.gov

Tissue Procurement: Miklos Palovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: LLAM10671 row: j column: 14
High quality sequence stop: 737.
Location/Qualifiers

FEATURES

source

1. 763
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792813"
/clone_1lb="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site: 1: BamHI; Site: 2: SalI-XhoI (gtcgag); Oligo-dt primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT

195 a 176 c 224 g 168 t

ORIGIN

Query Match 100.0%; Score 417; DB 12; Length 763;
Best Local Similarity 100.0%; Pred. No. 2.4e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGAGAGAGCAACCAATCTCCAGGCGAGCAATTTCTGTGTGAAGAATCAAGC 60
DB 295 GAGAGAGAGCAACCAATCTCCAGGCGAGCAATTTCTGTGTGAAGAATCAAGC 354
OY 61 AGGAGAAATTTGAGCGCGGAGATTGAGATTGAGAGCAAGACATGCTGCTGATT 120
DB 355 AGGAGAAATTTGAGCGCGGAGATTGAGATTGAGAGCAAGACATGCTGCTGATT 414
OY 121 CACTCAGGAAACGTGCTCAGGCGAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTA 180
DB 415 CACTCAGGAAACGTGCTCAGGCGAGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTA 474
OY 241 GCGGCTGCTGCTGTTGAACATCTACTCACTCAGATGAGATGAGTCACTGCTG 300
DB 535 GCGGCTGCTGCTGTTGAACATCTACTCACTCAGATGAGATGAGTCACTGCTG 594
OY 301 AGGCTGAGATTGAGAGTGTGCTTGAAGGCGAGAGTCAAGATGACTTCTGTGTGTA 360
DB 595 AGGCTGAGATTGAGAGTGTGCTTGAAGGCGAGAGTCAAGATGACTTCTGTGTGTA 654
OY 361 TTGACCGCTGTGTGAACATGATGATGGTGAGGCGCAACATGATCTGTGATGATGGG 417
DB 655 TTGACCGCTGTGTGAACATGATGATGGTGAGGCGCAACATGATCTGTGATGATGGG 711

RESULT 5

AU126038 764 bp mRNA linear EST 01-AUG-2002

LOCUS AU126038 764 bp mRNA linear EST 01-AUG-2002

DEFINITION AU126038 NT2RM4 Homo sapiens cDNA clone NT2RM4002613 5', mRNA sequence.

ACCESSION AU126038
VERSION AU126038.1 GI:10950754

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 764) Ota,T., Makamatsu,A., Ozawa,M., Ishii,S., Salto,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and Isogai,T.
TITLE	HRI human cDNA project (Ota,T., Makamatsu,A., Ozawa,M., Ishii,S., Salto,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S., Isogai,T.)
JOURNAL	Unpublished (2000)
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute. Location/Qualifiers 1..764 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="NT2RM4002613" /clone_11d="NT2RM4" /cell_type="teratocarcinoma" /cell_line="NT2" /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal precursor cells"
BASE COUNT	193 a 157 c 228 g 183 t 3 others
ORIGIN	
Query Match	100.0%; Score 417; DB 9; Length 764;
Best Local Similarity	100.0%; Prid. NO.2.4e-110;
Matches 417; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1	GAGAGACAGCAACCAACTCCAGAGGCGAGCAATTTCTGTGAGAAACATCAAGC 60
58	GAGAGACAGCAACCAACTCCAGAGGCGAGCAATTTCTGTGAGAAACATCAAGC 117
61	AGCGCAATTTGAGCGCGGAGATTGAGATTGCAGACCAACATCTGCTCTGATTT 120
118	AGCGCAATTTGAGCGCGGAGATTGAGATTGCAGACCAACATCTGCTCTGATTT 177
121	CACCTAGGAAAGCTCTAGAGGGGAGAAAGCCCTTGCTGCTTAAATAGTGGCTGTA 180
178	CACCTAGGAAAGCTCTAGAGGGGAGAAAGCCCTTGCTGCTTAAATAGTGGCTGTA 237
181	CACACATCAAGCCGAGAGCGGCTGATTGAGACACTGCGCCCTGGGGGCTCAGT 240
238	CACACATCAAGCCGAGAGCGGCTGATTGAGACACTGCGCCCTGGGGGCTCAGT 297
241	GCCCGTGTGCTTGTAACTACTACTCAACTCAGAAATGAAGTAGCTCAGACATGGCTG 300
298	GCCCGTGTGCTTGTAACTACTACTCAACTCAGAAATGAAGTAGCTCAGACATGGCTG 357
301	AGCGTGAAGTTCAGATGTTCCGTTGGAAAGGGCGAGTCAAGAGATACATTCTGGTGTGTA 360
358	AGCGTGAAGTTCAGATGTTCCGTTGGAAAGGGCGAGTCAAGAGATACATTCTGGTGTGTA 417
361	TTGACCGCTGTGTGAACATGATGGGTGGCGAGCCCAACATGATCTCTGATGATGGGG 417
418	TTGACCGCTGTGTGAACATGATGGGTGGCGAGCCCAACATGATCTCTGATGATGGGG 474

DEFINITION	602682515F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4815170 5', mRNA sequence.
ACCESSION	BC701625
VERSION	BC701625.1 GI:13972152
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 791)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10711 row: n column: 03 High quality sequence stop: 707. Location/Qualifiers
FEATURES	1..791
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4815170" /clone_1lb="NIH_MGC_95" /tissue_type="hippocampus" /lab_host="DH10B" /note="Organ: Brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgagag) ; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.5 kb and normalized to R0F 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NHGRI/NHRI, National Institutes of Health). Note: this is a NIH_MGC library."
BASE COUNT	196 a 193 c 226 g 176 t
ORIGIN	
Query Match	100.0%; Score 417; DB 12; Length 791;
Best Local Similarity	100.0%; Pred. No. 2.5e-110;
Matches 417; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1 GAGAGAGCAGCAAAACCACTCCAAAGGCGAGCAGCAGCAATTTCTGTGGAAGAACATCAAC	60
253 GAGAGAGCAGCAAAACCACTCCAAAGGCGAGCAGCAGCAATTTCTGTGGAAGAACATCAAC	311
61 AGGCAAGATTGGAGCGCGGAGATTGAGATTGCACAGCAAGCAGTCGTCTGATT	120
312 AGGCAAGATTGGAGCGCGGAGATTGAGATTGCACAGCAGCAGTCGTCTGATT	371
121 CACTAGAGAAAGCTGCTCAGGGGAGAAAGCCCTTGCTGCTTAATAATAGTGGCTGTA	180
372 CACTAGAGAAAGCTGCTCAGGGGAGAAAGCCCTTGCTGCTTAATAATAGTGGCTGTA	431
181 CACACATCAGACCCAGACAGCGGGTGTGATGAGCACTGTGTGCGCTGGGGCTCACT	240
432 CACACATCAGACCCAGACAGCGGGTGTGATGAGCACTGTGTGCGCTGGGGCTCACT	491
241 GCGCGTGTGCTCTGTATACATCTACTCAATCAACGAATGAAGTGAAGTGAAGCAGCGCG	300
492 GCGCGTGTGCTCTGTATACATCTACTCAACGAATGAAGTGAAGTGAAGCAGCGCG	551
301 AGCGTGAATTCAGTGTGCGTGTGGAAGGGGAGTCAAGAATGACTCTGTGTGTGTA	360
552 AGCGTGAATTCAGTGTGCGTGTGGAAGGGGAGTCAAGAATGACTCTGTGTGTGTA	611

Db 361 TTGACCGCTGTGTGACATGATGGTGGCGGCAACATGATCTGTGATGAGG 417
|||||
Db 612 TTGACCGCTGTGTGACATGATGGTGGCGGCAACATGATCTGTGATGAGG 668

RESULT 7
BM453920 867 bp mRNA linear EST:05-FEB-2002
LOCUS BM453920
DEFINITION AGENCOURT_6402958 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528995
ACCESSION BM453920
VERSION BM453920.1 GI:18502949
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 867)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps@email.nih.gov
Tissue Procurement: ATCC/DC/DP
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LML at:
<http://image.llnl.gov>
Plate: LLM12206 row: 1 column: 20
High quality sequence stop: 670.

FEATURES
Source Location/Qualifiers
1..867
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528995"
/clone_1lb="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI,
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 222 a 201 c 241 g 202 t 1 others
ORIGIN

Query Match 100.0%; Score 417; DB 13; Length 867;
Best Local Similarity 100.0%; Pred. No. 2.6e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GAGAGAGCAGCAACCACTCCAGAGGCGAGCAGCAATTTCTGTGTAAGAACATCAAGC 60
|||||
Db 202 GAGAGAGCAGCAACCACTCCAGAGGCGAGCAGCAATTTCTGTGTAAGAACATCAAGC 261

Db 61 AGGAGAAATTTGGAGCGCGGGAGATGAGATTGGAGAGCAACATGCTGCTGATTT 120
|||||
Db 262 AGGAGAAATTTGGAGCGCGGGAGATGAGATTGGAGAGCAACATGCTGCTGATTT 321

Db 121 CACTCAGAAAGCTGCTCAGAGGGGAGAGCCCTTGGTGTCTAAATAGTGGCTGTA 180
|||||
Db 322 CACTCAGAAAGCTGCTCAGAGGGGAGAGCCCTTGGTGTCTAAATAGTGGCTGTA 381

Db 181 CACACATCACAGCCAGAGCGGCTGTGATTGAGACACTGTGCTCCCTGGGGGCTCACT 240
|||||
Db 382 CACACATCACAGCCAGAGCGGCTGTGATTGAGACACTGTGCTCCCTGGGGGCTCACT 441

Db 241 GCCCGTGTCTGCTTGAATCTTACTCACTCAATCAATGAATGACTGAGACATGGCTG 300
|||||
Db 442 GCCCGTGTCTGCTTGAATCTTACTCACTCAATCAATGAATGACTGAGACATGGCTG 501

Db 301 AGCGTGAGTTGCAAGTTCCTTGGAGAGGCGAGTCAAGATGACTTCTGTGATGTA 360
|||||

Db 502 AGCGTGAGTTGCAAGTTCCTTGGAGAGGCGAGTCAAGATGACTTCTGTGATGTA 561
|||||
Db 361 TTGACCGCTGTGTGACATGATGGTGGCGGCAACATGATCTGTGATGAGG 417
|||||
Db 562 TTGACCGCTGTGTGACATGATGGTGGCGGCAACATGATCTGTGATGAGG 618

RESULT 8
AL551097 921 bp mRNA linear EST:16-FEB-2001
LOCUS AL551097
DEFINITION AL551097 LTI_NFL006_PL2 Homo sapiens cDNA clone CS01066VM07.5
ACCESSION AL551097
VERSION AL551097.1 GI:12888715
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 921)
L.W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Source Location/Qualifiers
1..921
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS01066VM07"
/clone_1lb="LTI_NFL006_PL2"
/tissue_type="Placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-Oligo(dt) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact: Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
Email: filang@life.com URL :
<http://fulllength.invitrogen.com>"

BASE COUNT 217 a 231 c 302 g 170 t 1 others
ORIGIN

Query Match 100.0%; Score 417; DB 9; Length 921;
Best Local Similarity 100.0%; Pred. No. 2.7e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GAGAGAGCAGCAACCACTCCAGAGGCGAGCAGCAATTTCTGTGTAAGAACATCAAGC 60
|||||
Db 461 GAGAGAGCAGCAACCACTCCAGAGGCGAGCAGCAATTTCTGTGTAAGAACATCAAGC 520

Db 61 AGGAGAAATTTGGAGCGCGGGAGATGAGATTGGAGAGCAACATGCTGCTGATTT 120
|||||
Db 521 AGGAGAAATTTGGAGCGCGGGAGATGAGATTGGAGAGCAACATGCTGCTGATTT 580

Db 121 CACTCAGAAAGCTGCTCAGAGGGGAGAGCCCTTGGTGTCTAAATAGTGGCTGTA 180
|||||
Db 581 CACTCAGAAAGCTGCTCAGAGGGGAGAGCCCTTGGTGTCTAAATAGTGGCTGTA 640

Db 641 CACACATCACAGCCAGAGCGGCTGTGATTGAGACACTGTGCTCCCTGGGGGCTCACT 700
|||||
Db 241 GCCCGTGTCTGCTTGAATCTTACTCACTCAATCAATGAATGACTGAGACATGGCTG 300
|||||
Db 701 GCCCGTGTCTGCTTGAATCTTACTCACTCAATCAATGAATGACTGAGACATGGCTG 760

Db 301 AGCGTGAGTTGCAAGTTCCTTGGAGAGGCGAGTCAAGATGACTTCTGTGATGTA 360
|||||

Db 761 AGCTGAGATTCACAGTGTTCCTTGGAAGGCGAGTCAGAGATGATCTGTGTGTGA 820
QY 361 TTGACCGCTGTGTGAACATGATGGTGGGCAACATGATCCGTGATGATGGG 417
|||||
Db 821 TTGACCGCTGTGTGAACATGATGGTGGGCAACATGATCCGTGATGATGGG 877

RESULT 9

BF033444 944 bp mRNA linear EST:20-OCT-2000
LOCUS BF033444
DEFINITION 601457970P1 NIH_MGC_66 Homo sapiens CDNA clone IMAGE:3861641 5',
mRNA sequence.
ACCESSION BF033444
VERSION BF033444.1 GI:10741156
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9598 row: k column: 18
High quality sequence stop: 600.
Location/Qualifiers
1. 944

FEATURES

Source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3861641"
/clone_lib="NIH_MGC_66"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: PCMV-SPORT6; Site: 1: Not;
Site 2: Salt; Cloned unidirectionally. Primer: oligo dt.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

BASE COUNT 244 a 198 c 273 g 229 t

ORIGIN

Query Match 100.0%; Score 417; DB 12; Length 944;
Best Local Similarity 100.0%; Pred. No. 2.7e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAGAGGCGACAGCAATTTCTGTGTGAAGACATCAAGC 60
Db 89 GAGAGAAGCAGCAACCACTCCAGAGGCGACAGCAATTTCTGTGTGAAGACATCAAGC 148
QY 61 AGGAGAGATTTGGAGCGCGGGAGATTGATTCAGAGCAAGACATGCTCTCTGATTT 120
Db 149 AGGAGAGATTTGGAGCGCGGGAGATTGATTCAGAGCAAGACATGCTCTCTGATTT 208
QY 121 CACTCAGGAAGAGCTGCTCAGAGGAGAGAGCCCTTGCTGTGCTAAATAGTGGGCTGTA 180
Db 209 CACTCAGGAAGAGCTGCTCAGAGGAGAGAGCCCTTGCTGTGCTAAATAGTGGGCTGTA 268
QY 181 CACACATCAGAGCCAGACAGCGGTGTTGATTGAGACACTGTGCCCTGGGGCTCAGT 240
Db 269 CACACATCAGAGCCAGACAGCGGTGTTGATTGAGACACTGTGCCCTGGGGCTCAGT 328
QY 241 GCCCGTGTGTGCTTGTAACTCTACTCAGATGGAAGTGAAGTGAAGTGAAGTGAAGTGA 300
Db 329 GCCCGTGTGTGCTTGTAACTCTACTCAGATGGAAGTGAAGTGAAGTGAAGTGAAGTGA 388

QY 301 AGCTGAGATTCACAGTGTTCCTTGGAAGGCGAGTCAGAGATGATCTGTGTGTGA 360
Db 389 AGCTGAGATTCACAGTGTTCCTTGGAAGGCGAGTCAGAGATGATCTGTGTGTGA 448
QY 361 TTGACCGCTGTGTGAACATGATGGTGGGCAACATGATCCGTGATGATGGG 417
|||||
Db 449 TTGACCGCTGTGTGAACATGATGGTGGGCAACATGATCCGTGATGATGGG 505

RESULT 10

BF259448 969 bp mRNA linear EST:13-FEB-2001
LOCUS BF259448
DEFINITION 602378544P1 NIH_MGC_92 Homo sapiens CDNA clone IMAGE:4509347 5',
mRNA sequence.
ACCESSION BF259448
VERSION BF259448.1 GI:12769264
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10389 row: g column: 12
High quality sequence stop: 700.
Location/Qualifiers
1. 969

FEATURES

Source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4509347"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: PCMV-SPORT6; Site: 1: Not;
Site 2: Salt; Cloned unidirectionally. oligo-dt primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."

BASE COUNT 234 a 219 c 290 g 226 t

ORIGIN

Query Match 100.0%; Score 417; DB 12; Length 969;
Best Local Similarity 100.0%; Pred. No. 2.7e-110;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGAAGCAGCAACCACTCCAGAGGCGACAGCAATTTCTGTGTGAAGACATCAAGC 60
Db 85 GAGAGAAGCAGCAACCACTCCAGAGGCGACAGCAATTTCTGTGTGAAGACATCAAGC 144
QY 61 AGGAGAGATTTGGAGCGCGGGAGATTGATTCAGAGCAAGACATGCTCTCTGATTT 120
Db 145 AGGAGAGATTTGGAGCGCGGGAGATTGATTCAGAGCAAGACATGCTCTCTGATTT 204
QY 121 CACTCAGGAAGAGCTGCTCAGAGGAGAGAGCCCTTGCTGTGCTAAATAGTGGGCTGTA 180
Db 205 CACTCAGGAAGAGCTGCTCAGAGGAGAGAGCCCTTGCTGTGCTAAATAGTGGGCTGTA 264
QY 181 CACACATCAGAGCCAGACAGCGGTGTTGATTGAGACACTGTGCCCTGGGGCTCAGT 240
Db 265 CACACATCAGAGCCAGACAGCGGTGTTGATTGAGACACTGTGCCCTGGGGCTCAGT 324
QY 241 GCCCGTGTGTGCTTGTAACTCTACTCAGATGGAAGTGAAGTGAAGTGAAGTGAAGTGA 300

DB 325 GCGCGTGTCTGCTTGAACATCTACCTCACTCACTGATGAGTGTGCTGACACTGCTG 384
QY 301 AGCGTGAAGTTGCAAGTGTGCTTGAAGGCGAGTCAGAGATGACTTCTGTGTGTGA 360
DB 385 AGCGTGAAGTTGCAAGTGTGCTTGAAGGCGAGTCAGAGATGACTTCTGTGTGTGA 444
QY 361 TTGACCGCTGTGTGAACATGATGAGTGTGCGAGCCCAACATGATCTGTGATGATGGG 417
DB 445 TTGACCGCTGTGTGAACATGATGAGTGTGCGAGCCCAACATGATCTGTGATGATGGG 501

RESULT 11
AL554850 858 bp mRNA 1linear EST_16-FEB-2001
LOCUS AL554850 LTI_NFL006.Pl2 Homo sapiens cDNA clone CS0D1087YL18 5
DEFINITION prime, mRNA sequence.
ACCESSION AL554850
VERSION AL554850.1 GI:12896025
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Source Location/Qualifiers
1..858
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1087YL18"
/clone_1bp="LTI_NFL006.Pl2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 212 a 206 c 255 g 182 t 3 others
ORIGIN

Query Match 99.9%; Score 416.6; DB 9; Length 858;
Best Local Similarity 99.8%; Pred. No. 3.4e-110;
Matches 416; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAAGCAGCAACCACTCCAGAGGCGACGACCAATTTCTGTGTGAAGAATCAACG 60
DB 296 GAGGAAGCAGCAACCACTCCAGAGGCGACGACCAATTTCTGTGTGAAGAATCAACG 355
QY 61 AGGAGAAATTTGGAGCGCGGGAGATTGATGAGAGCAAGACATGCTCTGTGATT 120
DB 356 AGGAGAAATTTGGAGCGCGGGAGATTGATGAGAGCAAGACATGCTCTGTGATT 415
QY 121 CACTCAGGAAACGTCTCAGGAGGAGAGAGCCCTTGGCTGTGTCTAAATAGTGGGCTGA 180
DB 416 CACTCAGGAAACGTCTCAGGAGGAGAGAGCCCTTGGCTGTGTCTAAATAGTGGGCTGA 475
QY 181 CACACATCAGAGCCAGACAGCGGTGTGATGAGACACTGTGTCCCTGGGGGCTCAGT 240
DB 476 CACACATCAGAGCCAGACAGCGGTGTGATGAGACACTGTGTCCCTGGGGGCTCAGT 535
QY 241 GCGCGTGTCTGCTTGAACATCTACCTCACTCACTGATGAGTGTGCTGACACTGCGGTG 300

DB 536 GCGCGTGTCTGCTTGAACATCTACCTCACTCACTGATGAGTGTGCTGACACTGCGGTG 595
QY 301 AGCGTGAAGTTGCAAGTGTGCTTGAAGGCGAGTCAGAGATGACTTCTGTGTGTGA 360
DB 596 AGCGTGAAGTTGCAAGTGTGCTTGAAGGCGAGTCAGAGATGACTTCTGTGTGTGA 655
QY 361 TTGACCGCTGTGTGAACATGATGAGTGTGCGAGCCCAACATGATCTGTGATGATGGG 417
DB 656 TTGACCGCTGTGTGAACATGATGAGTGTGCGAGCCCAACATGATCTGTGATGATGGG 712

RESULT 12
AL527928 938 bp mRNA 1linear EST_13-FEB-2001
LOCUS AL527928 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC027YC19 5
DEFINITION prime, mRNA sequence.
ACCESSION AL527928
VERSION AL527928.1 GI:12791421
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Source Location/Qualifiers
1..938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC027YC19"
/clone_1bp="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 241 a 208 c 275 g 212 t 2 others
ORIGIN

Query Match 99.9%; Score 416.6; DB 9; Length 938;
Best Local Similarity 99.8%; Pred. No. 3.5e-110;
Matches 416; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAAGCAGCAACCACTCCAGAGGCGACGACCAATTTCTGTGTGAAGAATCAACG 60
DB 241 GAGGAAGCAGCAACCACTCCAGAGGCGACGACCAATTTCTGTGTGAAGAATCAACG 300
QY 61 AGGAGAAATTTGGAGCGCGGGAGATTGATGAGAGCAAGACATGCTCTGTGATT 120
DB 301 AGGAGAAATTTGGAGCGCGGGAGATTGATGAGAGCAAGACATGCTCTGTGATT 360
QY 121 CACTCAGGAAACGTCTCAGGAGGAGAGCCCTTGGCTGTGTCTAAATAGTGGGCTGA 180
DB 361 CACTCAGGAAACGTCTCAGGAGGAGAGCCCTTGGCTGTGTCTAAATAGTGGGCTGA 420
QY 181 CACACATCAGAGCCAGACAGCGGTGTGATGAGACACTGTGTCCCTGGGGGCTCAGT 240
DB 421 CACACATCAGAGCCAGACAGCGGTGTGATGAGACACTGTGTCCCTGGGGGCTCAGT 480

QY 241 GCCCTGCTGCTTGTACATCTACTCAGATGAAGTAGCTGCACACTGCTG 300
|||||
DB 481 GCCCTGCTGCTTGTACATCTACTCAGATGAAGTAGCTGCACACTGCTG 540
QY 301 AGGCTGGAGTGCAGTGTTCCTGGAAGGGCGAGTCAGAGATGACTTCTGGTGTGA 360
|||||
DB 541 AGGCTGGAGTGCAGTGTTCCTGGAAGGGCGAGTCAGAGATGACTTCTGGTGTGA 600
QY 361 TTGACCGCTGTGTGAACATGATGGTGGCGAGGCCAACATGATCCTGGATGATGGG 417
|||||
DB 601 TTGACCGCTGTGTGAACATGATGGTGGCGAGGCCAACATGATCCTGGATGATGGG 657

RESULT 13
BG706693 701 bp mRNA linear: EST 07-MAY-2001
LOCUS 602671806F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4794744 5',
DEFINITION mRNA sequence.
ACCESSION BG706693
VERSION BG706693.1 GI:13982291
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 701)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LLM10676 row: k column: 01
High quality sequence stop: 701.

FEATURES
SOURCE
Location/Qualifiers
1..701
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4794744"
/clone_1lb="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTVVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 183 a 166 c 197 g 155 t
ORIGIN
Query Match 99.6%; Score 415.4; DB 12; Length 701;
Best Local Similarity 99.8%; Pred. No. 6.7e-110;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAGAACAGCAACCACTCCAGGCGAGCAGCAGCAATTCGTGTGAAGAACATCAAGC 60
|||||
DB 235 GAGAGAACAGCAACCACTCCAGGCGAGCAGCAGCAATTCGTGTGAAGAACATCAAGC 294
QY 61 AGGCAATTTGAGCGCGGAGATGAGATTGAGAGCAACATGCTGCTGATTT 120
|||||
DB 295 AGGCAATTTGAGCGCGGAGATGAGATTGAGAGCAACATGCTGCTGATTT 354

QY 121 CACTCAGGAACGTCGTCAGGGGAGAAAGCCCTGGCTGTCGTAATAATAGTGGCTGA 180
|||||
DB 355 CACTCAGGAACGTCGTCAGGGGAGAAAGCCCTGGCTGTCGTAATAATAGTGGCTGA 414
QY 181 CACACATCAGACCCAGACAGCGGCTGTTGATTGAGACACTGTGCCCTGGGGCTCAGT 240
|||||
DB 415 CACACATCAGACCCAGACAGCGGCTGTTGATTGAGACACTGTGCCCTGGGGCTCAGT 474
QY 241 GCCCTGCTGCTTGTGAACATCTACTCAGATGAAGTAGCTGCACACTGCTG 300
|||||
DB 475 GCCCTGCTGCTTGTGAACATCTACTCAGATGAAGTAGCTGCACACTGCTG 534
QY 301 AGGCTGGAGTGCAGTGTTCCTGGAAGGGCGAGTCAGAGATGACTTCTGGTGTGA 360
|||||
DB 535 AGGCTGGAGTGCAGTGTTCCTGGAAGGGCGAGTCAGAGATGACTTCTGGTGTGA 594
QY 361 TTGACCGCTGTGTGAACATGATGGTGGCGAGGCCAACATGATCCTGGATGATGGG 417
|||||
DB 595 TTGACCGCTGTGTGAACATGATGGTGGCGAGGCCAACATGATCCTGGATGATGGG 651

RESULT 14
B1668412 857 bp mRNA linear: EST 12-SEP-2001
LOCUS 603292688F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312133 5',
DEFINITION mRNA sequence.
ACCESSION B1668412
VERSION B1668412.1 GI:15582645
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 857)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LLM11790 row: p column: 22
High quality sequence stop: 755.

FEATURES
SOURCE
Location/Qualifiers
1..857
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5312133"
/clone_1lb="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTVVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 219 a 208 c 241 g 189 t
ORIGIN
Query Match 99.6%; Score 415.4; DB 13; Length 857;
Best Local Similarity 99.8%; Pred. No. 7.5e-110;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAGAGACACCAACCACTCAAGGAGACAGCAATTTCTGTGTGAAGAACATCAAGC 60
 DB 242 GAGAGAGACACCAACCACTCAAGGAGACAGCAATTTCTGTGTGAAGAACATCAAGC 301
 QY 61 AGGAGAGATTGGACGCCGGGAGATTGAGATTGACAGACACATGCTCTGTATT 120
 DB 302 AGGAGAGATTGGACGCCGGGAGATTGAGATTGACAGACACATGCTCTGTATT 361
 QY 121 CACTCAGAGAAAGCTGCTCAGGAGGAGAAAGCCCTTGCTGTGCTAAATAGTGGCTGTA 180
 DB 362 CACTCAGAGAAAGCTGCTCAGGAGGAGAAAGCCCTTGCTGTGCTAAATAGTGGCTGTA 421
 QY 181 CACACATCACAGCCCAAGACAGCGGTGTGATGAGACACTGTGTGCCCTGGGGCTCAAGT 240
 DB 422 CACACATCACAGCCCAAGACAGCGGTGTGATGAGACACTGTGTGCCCTGGGGCTCAAGT 481
 QY 241 GCCCGTGTCTGCTTGTAACTACTCACTCACTCAAGATGAGAGTCACTGCTGCTGCTG 300
 DB 482 GCCCGTGTCTGCTTGTAACTACTCACTCACTCAAGATGAGAGTCACTGCTGCTGCTG 541
 QY 301 AGGCTGAGATTGACAGTGTGCTTGTGAAGGCGAGTCAAGAAATGACTTGTGTGTGTA 360
 DB 542 AGGCTGAGATTGACAGTGTGCTTGTGAAGGCGAGTCAAGAAATGACTTGTGTGTGTA 601
 QY 361 TTGACCGCTGTGTAACATGATGATGGTGGGAGGCAACATATCTGTGATGATGGG 417
 DB 602 TTGACCGCTGTGTAACATGATGATGGTGGGAGGCAACATATCTGTGATGATGGG 658

RESULT 15
 BM449470 1042 bp MRNA linear EST 05-FEB-2002
 LOCUS BM449470

DEFINITION AGENCOURT_6400881 NIH_MGC_67 Homo sapiens CDNA clone IMAGE:5493677
 ACCESSION BM449470
 VERSION BM449470.1 GI:18498510
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1042)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ARCC

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov

Plate: LLM42117 row: m column: 06
 High quality sequence stop: 694.
 Location/Qualifiers
 1. 1042

FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5493677"
 /clone_lib="NIH_MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site: 1; NotI;
 Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

BASE COUNT 267 a 219 c 309 g 240 t 7 others
 ORIGIN

Query Match 99.3%; Score 414; DB 13; Length 1042;

Best Local Similarity 100.0%; Pred. No. 2,1e-109;
 Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 AGAAGACAGCAACCACTCAAGGAGACAGCAATTTCTGTGTGAAGAACATCAAGC 63
 DB 1 AGAAGACAGCAACCACTCAAGGAGACAGCAATTTCTGTGTGAAGAACATCAAGC 60
 QY 64 CAGAAATTGGACGCCGGGAGATTGAGATTGACAGACAAACATGCTCTGTATTGAC 123
 DB 61 CAGAAATTGGACGCCGGGAGATTGAGATTGACAGACAAACATGCTCTGTATTGAC 120
 QY 124 TCAGAGAAAGCTCAGGAGGAGAAAGCCCTTGCTGTGCTAAATAGTGGCTGTAC 183
 DB 121 TCAGAGAAAGCTCAGGAGGAGAAAGCCCTTGCTGTGCTAAATAGTGGCTGTAC 180
 QY 184 ACATCACAGCCCAAGACAGCGGTGTGATGAGACACTGTGTGCCCTGGGGCTCAAGT 243
 DB 181 ACATCACAGCCCAAGACAGCGGTGTGATGAGACACTGTGTGCCCTGGGGCTCAAGT 240
 QY 244 GCTGTCTGCTTGTAACTACTCACTCACTCAAGATGAAATGACTGCTGCTGAGG 303
 DB 241 GCTGTCTGCTTGTAACTACTCACTCACTCAAGATGAAATGACTGCTGCTGAGG 300
 QY 304 CTGAGATTGACAGTGTGCTTGTGAAGGCGAGTCAAGAAATGACTTGTGTGTGTA 363
 DB 301 CTGAGATTGACAGTGTGCTTGTGAAGGCGAGTCAAGAAATGACTTGTGTGTGTA 360
 QY 364 ACCGCTGTGTAAACATGATGATGGTGGCAGGCCAACATGATCTGTGATGATGGG 417
 DB 361 ACCGCTGTGTAAACATGATGATGGTGGCAGGCCAACATGATCTGTGATGATGGG 414

Search completed: April 21, 2003, 03:58:56
 Job time: 568.956 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 22:06:49 ; Search time 3345.48 Seconds
(without alignments)

11274.088 Million cell updates/sec

Title: US-09-782-051-1_COPY_549_1844

Perfect score: 1296

Sequence: 1 tccaagcagcagcagcattt.....aacctaattattacagatcac 1296

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 205640 segs, 14551402878 residues 4109280

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank

1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_ov.*
7: gb_ov.*
8: gb_ov.*
9: gb_ov.*
10: gb_ov.*
11: gb_ov.*
12: gb_ov.*
13: gb_ov.*
14: gb_ov.*
15: gb_ov.*
16: gb_ov.*
17: gb_ov.*
18: gb_ov.*
19: gb_ov.*
20: gb_ov.*
21: gb_ov.*
22: gb_ov.*
23: gb_ov.*
24: gb_ov.*
25: gb_ov.*
26: gb_ov.*
27: gb_ov.*
28: gb_ov.*
29: gb_ov.*
30: gb_ov.*
31: gb_ov.*
32: gb_ov.*
33: gb_ov.*
34: gb_ov.*
35: gb_ov.*
36: gb_ov.*
37: gb_ov.*
38: gb_ov.*
39: gb_ov.*
40: gb_ov.*
41: gb_ov.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1296	100.0	2258	9	HSU82761	U82761 Homo sapien
2	1296	100.0	2526	9	BC007576	BC007576 Homo sapi
3	1296	100.0	2552	9	BC016942	BC016942 Homo sapi
4	1296	100.0	2563	6	AX029176	AX029176 Sequence
5	1296	100.0	2583	9	BC010681	BC010681 Homo sapi
6	1296	100.0	2577	9	AF315687	AF315687 Homo sapi
7	1294.4	99.9	2510	9	HSU800298	AL049954 Homo sapi
8	1166.4	90.0	1772	10	BC018218	BC018218 Mus muscu
9	863	66.6	1987	9	AK025372	AK025372 Homo sapi
10	861.4	66.5	2030	9	BC008349	BC008349 Homo sapi
11	861.4	66.5	5025	9	AB020635	AB020635 Homo sapi
12	861.4	66.5	5052	9	BC024325	BC024325 Homo sapi
13	644	49.7	1369	9	AF035319	AF035319 Homo sapi
14	631	48.7	1792	3	AY113501	AY113501 Drosophila
15	599	46.2	1323	9	BC003631	BC003631 Homo sapi
16	486.2	37.5	40320	2	AC015387	AC015387 Drosophila
17	486.2	37.5	172372	3	AC010110	AC010110 Drosophila
18	412.6	31.8	2057	10	MUSSAHR	L32836 Mus musculu
19	411	31.7	2067	10	BC015304	BC015304 Mus muscu
20	411	31.7	190612	2	AC129093	AC129093 Mus muscu
21	406.6	31.4	52423	2	AC020395	AC020395 Drosophila
22	406.6	31.4	80423	3	DROABDB	L07835 Drosophila
23	406.6	31.4	175335	3	AC091636	AC091636 Drosophila
24	406.6	31.4	223098	3	AE003715	AE003715 Drosophila
25	406.6	31.4	338234	3	DMU31961	U31961 Drosophila
26	406.6	31.3	1302	6	AX458351	AX458351 Sequence
27	402.4	31.0	2084	9	HUMAHCY2	M61832 Human S-ade
28	402.4	31.0	2175	9	BC010018	BC010018 Homo sapi
29	402.4	31.0	2183	9	BC011606	BC011606 Homo sapi
30	402.4	31.0	2355	9	AK097610	AK097610 Homo sapi
31	401.4	31.0	2029	10	RATAHNA	M15185 Rat S-adeno
32	399	30.8	2088	5	XLAT7835	AJ007835 Xenopus L
33	396.6	30.6	5050	3	DMBX200	X13168 Drosophila
34	394.2	30.4	2129	5	XELAHN	L35559 Xenopus lae
35	375	28.9	1603	3	AY102668	AY102668 Drosophila
36	373.8	28.8	1564	3	DMARCTGEN	X95836 D.melanogas
37	359	27.7	1617	3	AF080546	AF080546 Anopheles
38	354	27.3	1631	3	LEISDSH	M76556 L.donovani
39	346	26.7	1315	3	DDIAHNA	M19937 S.lime mold
40	337	26.0	384	6	AX407699	AX407699 Sequence
41	334.4	25.8	173839	2	AC118421	AC118421 Rattus no
42	322.4	24.9	37187	8	AB004537	AB004537 Schizosac
43	322.4	24.9	43757	8	SPBC8D2	AL022072 S.pombe C
44	320.8	24.8	128638	2	AC120808	AC120808 Rattus no
45	314.4	24.3	2211	6	AR144978	AR144978 Sequence

ALIGNMENTS

RESULT 1

LOCUS HSD82761 2258 bp mRNA linear PRI 17-DEC-2001

DEFINITION Homo sapiens S-adenosyl homocysteine hydrolase homolog (XPVKona)

ACCESSION U82761

VERSION U82761

KEYWORDS mRNA, complete cds.

SOURCE U82761.1 GI:2852124

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2258)

Volpe,J.F.G., Busch,D.B., Coleman,D.W., Ziffer,D.W., Yu,Y.,

Pred. No. is the number of results predicted by chance to have a

TITLE Nagasawa, H. and Little, J.B.
 Increased ultraviolet sensitivity and chromosomal instability related to P53 function in the xeroderma pigmentosum variant
JOURNAL Cancer Res. 59 (5), 1102-1108 (1999)
MEDLINE 99168517
PUBMED 10070969
REFERENCE 2 (bases 1 to 2258)
AUTHORS Volpe, J.P.G., McDowell, M., Jostes, R.F., Afzal, V., Sadinski, M., Trask, B.J., Legerkl, R. and Cleaver, J.E.
TITLE Complementation of chromosomal instability in the xeroderma pigmentosum variant by a gene on human chromosome 1 with homology to S-adenosyl homocysteine hydrolase
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2258)
AUTHORS Volpe, J.P.G., McDowell, M. and Cleaver, J.E.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1996) Dermatology, UCSF, 3rd and Parnassus, Box 0750, San Francisco, CA 94143, USA
FEATURES Location/Qualifiers
source 1. 2258
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /cell_type="fibroblasts"
 /issue_type="skin"
 1. 2258
 /gene="XPV-kona"
 48. 1550
 /gene="XPV-kona"
 /note="xeroderma pigmentosum variant"
 /codon_start=-1
 /product="S-adenosyl homocysteine hydrolase homolog"
 /protein_id="AAC01960.1"
 /db_xref="GI:2852125"
 /translation="MATVTRKPKKQIOFADMDQETKPKTKGRSISRSISOSSTDS
 YSSASITBSSDDEVSPREKQNTNKGSSNFKVKNIKQAFGRRELEIAQDMSALIS
 LRRKQAEKPLAKIKIVGCTHTIAOTAVLLETICAGACQRMWACNITSTONFAAL
 AEAQVAFANKGESEDFWVICIDRCVMDQMANMIDGDDTHWYKXPKYFKFI
 RGVESVTVGHRLYOLSKAKLCVPMANVDSVTKOKPNLCCRSIIDLKRTTD
 VMGKQVYVCGEYGVKGCCKALKAIGATVITTEIDPICALQMDGFEVVLNFI
 ROVDVITTCGNKNNVTRHLDPMKNSCTYCNMGSHTETIDVTSLRPELTWERVRSO
 VDIWVMDGRVLLAEGRLNLISCTVPFVLSITATQALALIELYNPBERKXKOD
 VYLLPKMEYVSLHLPSEDAHLTELTDQAKYLLGINKNGPKPNYRYR"
BASE COUNT 603 a 513 c 542 g 600 t
ORIGIN
 Query Match 100.0%; Score 1296; DB 9; Length 2258;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 361 GATGGGTGGCAGGCCAACATGATCTCGATGATGGGGAGACTTAACCATGCGTTAT 420
 Db 612 GATGGGTGGCAGGCCAACATGATCTCGATGATGGGGAGACTTAACCATGCGTTAT 671
 QY 421 AAGAATATCCAAACGTTTAAAGATCCGAGGCAATTTGGAGAGAGCGTGAAGT 480
 Db 672 AAGAATATCCAAACGTTTAAAGATCCGAGGCAATTTGGAGAGAGCGTGAAGT 731
 QY 481 GTTACAGGCGTGTACAGCTCTCAAGCTGGGAAGCTCTGTTCGCGCATGAGTC 540
 Db 732 GTTACAGGCGTGTACAGCTCTCAAGCTGGGAAGCTCTGTTCGCGCATGAGTC 791
 QY 541 AATGATCTGTACCAACAGAGTTGATTAATCTGTACTCTGCGGAGAAATTCATTG 600
 Db 792 AATGATCTGTACCAACAGAGTTGATTAATCTGTACTCTGCGGAGAAATTCATTG 851
 QY 601 GATGGCCTGAAGAGAGACACAGATGTGATGTTGGTGGGAACAAGTGTGTGTGGC 660
 Db 852 GATGGCCTGAAGAGAGACACAGATGTGATGTTGGTGGGAACAAGTGTGTGTGGC 911
 QY 661 TATGTAGTAGGCAAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
 Db 912 TATGTAGTAGGCAAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 971
 QY 721 ATTACCGAATCGAAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
 Db 972 ATTACCGAATCGAAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1031
 QY 781 AAGCTAAATGAAGTATCCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
 Db 1032 AAGCTAAATGAAGTATCCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1091
 QY 841 GTATGACACGCGGAGCACTTGGATTCGATGAAAACATTTATTCGTATGCAATGGGC 900
 Db 1092 GTATGACACGCGGAGCACTTGGATTCGATGAAAACATTTATTCGTATGCAATGGGC 1151
 QY 901 CACTCCAAACAGAAATCGATGTGACACAGCTCCGCACTCCGAGCTAGCTGGAGCGA 960
 Db 1152 CACTCCAAACAGAAATCGATGTGACACAGCTCCGCACTCCGAGCTAGCTGGAGCGA 1211
 QY 961 GTACGTTCTCAGGTGAGCATGTATCTGCGCAGATGGCAACGATGTCTCTGGCA 1020
 Db 1212 GTACGTTCTCAGGTGAGCATGTATCTGCGCAGATGGCAACGATGTCTCTGGCA 1271
 QY 1021 GAGGTGTCTACATCAATTTAGCTGTCCACAGTTCCTCACTTGTCTCATCA 1080
 Db 1272 GAGGTGTCTACATCAATTTAGCTGTCCACAGTTCCTCACTTGTCTCATCA 1331
 QY 1081 GCCACACACAGGCTTGGCACTGATGAACTGATATGACCGAGGCGGATACAG 1140
 Db 1332 GCCACACACAGGCTTGGCACTGATGAACTGATATGACCGAGGCGGATACAG 1391
 QY 1141 CAGATGTGTACTGTCTCTCAAGAAATGATGATAGTTCGACAGCTTCATCTGCA 1200
 Db 1392 CAGATGTGTACTGTCTCTCAAGAAATGATGATAGTTCGACAGCTTCATCTGCA 1451
 QY 1201 TCATTTGATGCCACCTTACAGACTGACAGATGACCAAGAAATATCTGGACTCAAC 1260
 Db 1452 TCATTTGATGCCACCTTACAGACTGACAGATGACCAAGAAATATCTGGACTCAAC 1511
 QY 1261 AAAAATGGCCATTCAAACTTAATTTATTCAGATAC 1296
 Db 1512 AAAAATGGCCATTCAAACTTAATTTATTCAGATAC 1547
RESULT 2.
LOCUS BC007576 2526 bp. mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
ACCESSION BC007576
VERSION BC007576.1 GI:14043176

QY 1201 TCATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAAC 1260
 DB 1712 TCATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAAC 1771
 QY 1261 AAAATGGGCGCATTCACCACTATATTTACAGATAC 1296
 DB 1772 AAAATGGGCGCATTCACCACTATATTTACAGATAC 1807

RESULT 3

LOCUS BC016942 2552 bp mRNA linear PRI-09-NOV-2001
 DEFINITION Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone

ACCESSION BC016942
 VERSION BC016942.1 GI:16877386
 KEYWORDS MGC.

SOURCE Homo sapiens.
 ORGANISM Homo sapiens, Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2552)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (05-NOV-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpax11.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
 Series: IRAX Plate: 20 Row: a Column: 15
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein.

FEATURES

SOURCE

Location/Qualifiers
 1..2552
 /organism="Homo sapiens"
 /db_xref="locusid:10768"
 /db_xref="taxon:9606"
 /clone="MGC:21453 IMAGE:3450568"
 /tissue_type="Placenta, choriochorionoma"
 /clone_id="NH_MGC_10"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 338..1840
 /codon_start=1
 /product="S-adenosylhomocysteine hydrolase-like 1"
 /protein_id="AAH16942.1"
 /db_xref="GI:16877387"
 /translation="MATVTAKPKKIOFADDMOEFKEPTKGRSLRSISOSTDS
 YSSASYSDDSDDESPREKOOTNSKSGCNVKNIKIAEGRREIEIAEDMSLIS
 LKRAOGEKPLAGAKIVGCTHITNOTATLITLICALACGMSACNISTONEVAAL
 AAGVAYPAFMGSEDEDFWNCIDRCVNDGQANILDDGDLTHVYKTKPPNPKKI
 KGIYESVYGVHRLYQLSKAGALCVPANVNDVTKQKFDNLKCCRESILDGKLTETD
 VMFGKQYVGVGEYEVGCGCAALKAALCAIYIIEIDICALQCMDFRVLKLTETD
 RDYVIVICTGNKNVYREHIDRMKNSCIYCNMGSTNCTIATSLRPELWERYRSQ
 VDYIVPDKREVLLAEGRILNDSCTVPTLVLSSTATTOALILELNAEGRVKOD
 VYLLPKRDEVVAISLHLPSPAHLELDDAKYIGLNKNPFRNRYR"

CDS

BASE COUNT

644 a 600 c 671 g 637 t

ORIGIN

Query Match 100.0%; Score 1296; DB 9; Length 2552;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCAAAGGAGAGCAAGCAATTCCTGTGTGAAGAACATCAACAGCAGCAATTTGGACCCGG 60
 DB 542 TCCAAAGGAGAGCAAGCAATTCCTGTGTGAAGAACATCAACAGCAGCAATTTGGACCCGG 601
 QY 61 GAGATTGAGATTGACAGAGCAAGCAATTCCTGTGTGAAGAACATTCAGCAAGAGTGTCTAG 120
 DB 602 GAGATTGAGATTGACAGAGCAAGCAATTCCTGTGTGAAGAACATTCAGCAAGAGTGTCTAG 661
 QY 121 GGGGAGAACCCCTTGGCTGTGTCTAAATATGTGGCTGTACACATCAACAGCCGAGACA 180
 DB 662 GGGGAGAACCCCTTGGCTGTGTCTAAATATGTGGCTGTACACATCAACAGCCGAGACA 721
 QY 181 GCGGCTGTGTGTGAGACATCTGTGCGCCGGGGGCGTCAATGCGCGGTGTGTATAC 240
 DB 722 GCGGCTGTGTGTGAGACATCTGTGCGCCGGGGGCGTCAATGCGCGGTGTGTATAC 781
 QY 241 ATCTACTCACTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 DB 782 ATCTACTCACTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 841
 QY 301 GCTTGAAGAGGCGAGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 DB 842 GCTTGAAGAGGCGAGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 901
 QY 361 GATGGGTGCGAGGCGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 DB 902 GATGGGTGCGAGGCGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961
 QY 421 AAGAATATCAACAGCTGTTTAAGATCCAGAGCATTTGGAGAGAGCGTGAATGAT 480
 DB 962 AAGAATATCAACAGCTGTTTAAGATCCAGAGCATTTGGAGAGAGCGTGAATGAT 1021
 QY 481 GTTCACAGGCTGTATCAGCTTCTCCAAAGCTGGAGAGCTGTGTGCTCGGCATGAACGTC 540
 DB 1022 GTTCACAGGCTGTATCAGCTTCTCCAAAGCTGGAGAGCTGTGTGCTCGGCATGAACGTC 1081
 QY 541 AATGATTCCTGTATCCAAAG 600
 DB 1082 AATGATTCCTGTATCCAAAG 1141
 QY 601 GATGCGCTGAAG 660
 DB 1142 GATGCGCTGAAG 1201
 QY 661 TATGCTGAGGTAGAGCAAGGCTGTGTGCTGTCTCAAAAGCTTTGGAGCAATTTGTCTAC 720
 DB 1202 TATGCTGAGGTAGAGCAAGGCTGTGTGCTGTCTCAAAAGCTTTGGAGCAATTTGTCTAC 1261
 QY 721 ATTACCGAATATGACCCCATCTGTCTGTGACAGGCTGTGATGATGATGATGATGATGATG 780
 DB 1262 ATTACCGAATATGACCCCATCTGTGTGTGACAGGCTGTGATGATGATGATGATGATGATG 1321
 QY 781 AAGCTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 DB 1322 AAGCTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1381
 QY 841 GTAGTGACACGAGGAGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 DB 1382 GTAGTGACACGAGGAGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1441
 QY 901 CACTCCACACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 DB 1442 CACTCCACACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1501
 QY 961 GTACGTTCCAGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 DB 1502 GTACGTTCCAGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1561

Oy	1021	GAGGGTCGTCTCTCAATTTGAGTGCATCAAGTCCACCTTTGTGTGTCATACACA	1080
Db	1563	GAGGGTCGTCTCAATTTGAGTGCATCAAGTCCACCTTTGTGTGTCATACACA	1621
Oy	1081	GCCACACACAGGGCTTTGGCACTGATAGAACTCTATTAATGCACCCGAGGGGGCATACAG	1140
Db	1622	GCCACACACAGGGCTTTGGCACTGATAGAACTCTATTAATGCACCCGAGGGGGCATACAG	1681
Oy	1141	CAGGATGTGTACTTCTCTCTCTCAAGAAAATGATGTAATCGTGCACAGCTTGATCTGCCA	1200
Db	1682	CAGGATGTGTACTTCTCTCTCTCAAGAAAATGATGTAATCGTGCACAGCTTGATCTGCCA	1741
Oy	1201	TCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAAC	1260
Db	1742	TCATTTGATGCCCACTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAAC	1801
Oy	1261	AAAAATGGGCCATTCAAAACCTTAATATATACAAATAC	1296
Db	1802	AAAAATGGGCCATTCAAAACCTTAATATATACAAATAC	1837

[illegible]

REFERENCE	1 (bases 1 to 2563)
AUTHORS	Hart, D.N.
TITLE	Enzyme having s-adenosyl-l-homocysteine hydrolase (ancy) type activity
JOURNAL	Patent: WO 9814562-A, 1 09-APR-1998;

FEATURES	LOCATION/Qualifiers
SOURCE	1. 2563
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
CDS	<3. 1847

the full 5' nucleotide sequence. The initiation codon has yet to be identified."

/codon_start=1
/protein_id="CAC09285.1"
/db_xref="gi:10190061"
/translation="RGVVGARSCTCFMPSGHRRCRCPALGCTDKRWAKATAPKQPQLD
GAGRVADNRVSECGAARAGRAPESBERGGGGSAGARAGGSMSPDMLPVEGELK
AKLEDEDKRTSMFATVITKAPKQIQFADDMQELTFKPTKRRSLISRSISOSTSDYS
SAAASYTSDSDVESVPRERKQNTNSGSSNFCFKEAGGRREIIEADMSALISL
KRAGEKPLAGAKIVGCTHITACAVTILETICALAGRWACNAVYISQNSVVAALAE
AGVAFPMKGESEDDFMWCCIDRCVNMAGGMLLDDGSDLTWHNYCKPWFKKIR
EVESTVWVHGYLOLSKAGKLCVPAMNVNSETDPCALNCCESITDGVKLTIDVW
EGGSOVVGCGVGEYGGKCCALIKALIVYITETDPCALNCCESITDGVKLTIDVW
VDVYITDCKNKNTVTEGRILBMRKSCVYVNGNHNTEIDYSLRTPELTWRVYSOUT
HVTPDCKRVYVALLAGRLNLSCSTVPEFVSLITATDALLIELTNBESRTIKODVVA
LLPKMKDEYVALLSPSDAHLETLTDQAKYIDLNNKGPMPKPYR"

646 a 604 c 677 g 636 t

	Query Match	Score 1296:	DB 6:	Length 2563:
	Best Local Similarity	100.0%	Pred. No. 0:	
	Matches 1296:	Conservative 0:	Mismatches 0:	Indels 0: Gaps 0:
QY	1	TCCAAGGGCAGACGAAATTCCTGCTGTTGAAGAACATCAACAGACAGACGAAATTTGGACGCCG	60	
DB	549	TCCAAGGGCAGACGAAATTCCTGCTGTTGAAGAACATCAACAGACGACGAAATTTGGACGCCG	608	
QY	61	GAGATTGAGATTGCAGACGACGAAACATGCTGCTGTTGAATTTCACTCAGGAACGTCGTCAG	120	

D	609	GAGATGTGATTTGGACAGCAAGACATGTCTGCTCATTTCACTGACGAAAGCTGTACG	668
Q	121	GGGGGAAGACCCCTTGGCTGCTTAAATATGTGGCTGTACACACATCACACCCACACA	180
D	669	GGGGGAAGACCCCTTGGCTGCTTAAATATGTGGCTGTACACACATCACACCCACACA	728
Q	181	GGGGTGTGATTTGAGACACTGTGGCCCTGGGGGGGCACTAGCGCGCTGTCTGTAAAC	240
D	729	GGGGTGTGATTTGAGACACTGTGGCCCTGGGGGGGCACTAGCGCGCTGTCTGTAAAC	788
Q	241	ATCTACTCACTCAGAAATGAATAGTAGCTGCACACTGGCTGAGGCTGGAGTTGCAGTGTTC	300
D	789	ATCTACTCACTCAGAAATGAATAGTAGCTGCACACTGGCTGAGGCTGGAGTTGCAGTGTTC	848
Q	301	GCTTGGAAAGGCCGAGTCAGAAAGATGACTTCTGTGTGTATTGTAGCCGCTGTGTAAATG	360
D	849	GCTTGGAAAGGCCGAGTCAGAAAGATGACTTCTGTGTGTATTGTAGCCGCTGTGTAAATG	908
Q	361	GATGGGTGGCAAGGCCAACAATGATCCCTGGATGATGGGGGAGACTTAAACCAACGGGTTAT	420
D	909	GATGGGTGGCAAGGCCAACAATGATCCCTGGATGATGGGGGAGACTTAAACCAACGGGTTAT	968
Q	421	AAGAAGTATCCAAACGTTTAAAGAAATCCAGAGCAATTGTGGAGAGAGCGTGCATGCT	480
D	969	AAGAAGTATCCAAACGTTTAAAGAAATCCAGAGCAATTGTGGAGAGAGCGTGCATGCT	1028
Q	481	GTTCAAGCTGTATTCAGCTCTCCAAAGCTGGGAAGCTGTGTGTTCCGGCATGAAGCTTC	540
D	1029	GTTCAAGCTGTATTCAGCTCTCCAAAGCTGGGAAGCTGTGTGTTCCGGCATGAAGCTTC	1088
Q	541	AATGATTTCTTTACCAACAGAAATTTGATTAATCTTACTGCTGCCGAGATTCATTTTG	600
D	1089	AATGATTTCTTTACCAACAGAAATTTGATTAATCTTACTGCTGCCGAGATTCATTTTG	1148
Q	601	GATGACCTGAAGAGAGACACAGATGTGATTTTGTGTGGGAAACAATGGTGGTGTGATGC	660
D	1149	GATGACCTGAAGAGAGACACAGATGTGATTTTGTGTGGGAAACAATGGTGGTGTGATGC	1208
Q	661	TATGGTAGGATGAGCAAGGGCTGCTGTGCTCTCTCAAAAGCTTTTGGAGCAATTTGTAC	720
D	1209	TATGGTAGGATGAGCAAGGGCTGCTGTGCTCTCTCAAAAGCTTTTGGAGCAATTTGTAC	1268
Q	721	ATTACCGAAATCGACCCCATCTGTGCTCTCAGACCTCGATGGATGGGTTCAGGGTGTGA	780
D	1269	ATTACCGAAATCGACCCCATCTGTGCTCTCAGACCTCGATGGATGGGTTCAGGGTGTGA	1328
Q	781	AAGCTAATGAAGTATCCCGGCAAGTGTGATGTGTAATTAATCTGCACAGAGAAATAGAT	840
D	1329	AAGCTAATGAAGTATCCCGGCAAGTGTGATGTGTAATTAATCTGCACAGAGAAATAGAT	1388
Q	841	GTAGTGACACGGGAGCACTTGGATCGATTAATAAACAGTTGTATCGTATGCAATATGGGC	900
D	1389	GTAGTGACACGGGAGCACTTGGATCGATTAATAAACAGTTGTATCGTATGCAATATGGGC	1448
Q	901	CACCTCAACACAGAAATCGATGTGACCAAGCTCCGACCTCCGAGCTGACGTGGAGCGGA	960
D	1449	CACCTCAACACAGAAATCGATGTGACCAAGCTCCGACCTCCGAGCTGACGTGGAGCGGA	1508
Q	961	GTACGTTCTCAGGTGGACATGTCTATCTGCGCAGATGGGAAACAGATGTCTCTCTGGCA	1020
D	1509	GTACGTTCTCAGGTGGACATGTCTATCTGCGCAGATGGGAAACAGATGTCTCTCTGGCA	1568
Q	1021	GAGGCTGCTATCAATTTTGTAGCTGTCACAGTTCCACACTTGTGTTCTGTCATTCACA	1080
D	1569	GAGGCTGCTATCAATTTTGTAGCTGTCACAGTTCCACACTTGTGTTCTGTCATTCACA	1628
Q	1081	GCCACACACAGGCTTTGGCACTGATGAACCTGTATATGACCCGAGGGGCGATACAG	1140
D	1629	GCCACACACAGGCTTTGGCACTGATGAACCTGTATATGACCCGAGGGGCGATACAG	1688
Q	1141	CAGGATGTGTACTTGTCTCTTAAGAAATAGATTAAGTTGCCAGCTGTGATCTGGCA	1200

DB	1689	CAGAGTGTACTCTTCTCTCCTAGAAAAATGGATGAATACGTTGGCCAGCTGCATCTGGCA	1748
QY	1201	TCATTTGATGGCCACCTTACAGAGCTGACGACATATGACCAAGCAAAATATCTGGAGATCAAC	1260
DB	1749	TCATTTGATGGCCACCTTACAGAGCTGACGATGACCAAGCAAAATATCTGGAGATCAAC	1808
QY	1261	AAAAATGGGCCATTCAAACCTTATTTATTCAGATAC	1296
DB	1809	AAAAATGGGCCATTCAAACCTTATTTATTCAGATAC	1844

RESULT 5	
BC010681	
LOCUS	
DEFINITION	
BC010681	2583 bp mRNA linear PRI 12-JUL-2001
Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone	
MGc:8936 IMAGE:3853747, mRNA, complete cds.	

VERSION	BC010681.1	GI:14715037
---------	------------	-------------

SOURCE Homo sapiens.

ORGANIS

REFERENCE 1 (bases 1 to 2583)

JOURNAL

REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT	Contact: MGC help desk

submitted (10-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapds-rt@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: vallalona@bcm.tmc.edu

Vallalona, D.K., Luna, R.A., Hale, S.M., Huliyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Auzny, D.M., Gibbs, R.A.

FEATURES

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 13 Row: C Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5729723.

Location/Qualifiers

CDS

```

1. 25833
/organism="Homo sapiens"
/db_xref="locusid:10768"
/db_xref="taxon:9606"
/clone="MGC:8936 IMAGE:3853747"
/tissue_type="Colon, adenocarcinoma"
/clone_1ib="NH_MGC_65"
/lab_host="DH10B"
/notes="vector: pcMV-SpORe6"
367. 1869
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like 1"
/protein_id="AAH10681.1"
/db_xref="GI:14715038"
/translation="MAVYTAAPKAKQIQFADDDQEFKFKTKYGRSLSSKSSISQSDTDS
ISSASTSDSDSDVSPREKQIQNSGSSSFFCYKSLKQAEFRRLRELEADQMSALIS
LKRKAQOEKPLAGKATVGCCHTIAAQTVALETCALGACRMSACNETYSTONFAALAA
AEAAVAFAMGSEEDPFWMCICVVMGMOAMNLLDGGDLTHVYKKAQNPYK
RGIVSESTVGPHRLTSGKGLCPVAMANNQSVTKORFDNIYCRSESLDGLKRTTET
VMEGKQVYVNGGSEBNGCCGCAALKALGAYVITTEIDPCALQACMDGRVYKLAEYV
KQADVIVTGKNNKVAIVREHLDNRKNSCTVCAVNGNSHTEDVTSLEPLTAERGRKQD
KQADVIMPEGRKVVIVLAEGRNLNCSVTEPLVSTAVTQALALELNAAPGRKQD

```

YVLLPKKNDYVASLHLPSEDAHTEITDDQAKYLGINKNGPFKNRYR"
BASE COUNT 648 a 606 c 688 g 641 t
ORIGIN

Query Match	100.0%;	Score 1296;	DB 9;	Length 2583;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1296;	Conservative 0;	Mismatches	0;	Indels 0;
				Gaps 0;

OY	1	TCGAAGGGACACACATTTTCTGTGTGAACAATCAACAGGACAAATTTTGACGGG	60
Db	571	TCGAAGGGACACACATTTTCTGTGTGAACAATCAACAGGACAAATTTTGACGGG	630
OY	61	GAGATTGAGATTGACAGACAAAGACATGCTGCTCTGATTTCACTCAGAAAGTGTGAG	120
Db	631	GAGATTGAGATTGACAGACAAAGACATGCTGCTCTGATTTCACTCAGAAAGTGTGAG	690
OY	121	GGGGAGAAAGCCCTTGGCTGCTAAATATGTGGGCTGTACACATCAACAGCCACACA	180
Db	691	GGGGAGAAAGCCCTTGGCTGCTAAATATGTGGGCTGTACACATCAACAGCCACACA	750
OY	181	GGGGTGTGATTGAGACACTGTGGCCCTGAGGGGCTCAGTGGCCGCTGCTGTTAAC	240
Db	751	GGGGTGTGATTGAGACACTGTGGCCCTGAGGGGCTCAGTGGCCGCTGCTGTTAAC	810
OY	241	ATCTACTCACTCAGAAATGAAGTAGCTGACAGCACTGGCTGAGGCTGGAAGTTCAGTTC	300
Db	811	ATCTACTCACTCAGAAATGAAGTAGCTGACAGCACTGGCTGAGGCTGGAAGTTCAGTTC	870
OY	301	GCTTGGAAAGGGCGAGTCAAGAAAGATGATTTGTGTGTGTATTTACCCGCTGTGAAACAG	360
Db	871	GCTTGGAAAGGGCGAGTCAAGAAAGATGATTTGTGTGTGTATTTACCCGCTGTGAAACAG	930
OY	361	GATGGGTGGCAGGCGCAACATGATCCTGGATGATGGGGAGACTTAACCACTGGTTTAT	420
Db	931	GATGGGTGGCAGGCGCAACATGATCCTGGATGATGGGGAGACTTAACCACTGGTTTAT	990
OY	421	AAGAAATATCCAAACGTGTTTAAGAAAGATCCGAGCAATTGTGAGAGAGAGCGTAGCGT	480
Db	991	AAGAAATATCCAAACGTGTTTAAGAAAGATCCGAGCAATTGTGAGAGAGAGCGTAGCGT	1050
OY	481	GTTCAAGGGCTGTACAGCTCTCCAAAGCTGGGGAAGCTGTGTCTCCGGCCATGAAAGCTC	540
Db	1051	GTTCAAGGGCTGTACAGCTCTCCAAAGCTGGGGAAGCTGTGTCTCCGGCCATGAAAGCTC	1110
OY	541	AATGATTTCTGTTACCAACAGAAATTTGATAATTTACTAGCTGCCGAGAAATCCATTTTG	600
Db	1111	AATGATTTCTGTTACCAACAGAAATTTGATAATTTACTAGCTGCCGAGAAATCCATTTTG	1170
OY	601	GATGGCCTGAAGAGAGACCACAGATGTGATGTTTGTGTGGGAACAAGTGTGTGTGTGC	660
Db	1171	GATGGCCTGAAGAGAGACCACAGATGTGATGTTTGTGTGGGAACAAGTGTGTGTGTGC	1230
OY	661	TATGTGAGTAGGACAGAGGCGTGTGTGCGTCCTCAAGCTCTTGAGACATTTGTCTAC	720
Db	1231	TATGTGAGTAGGACAGAGGCGTGTGTGCGTCCTCAAGCTCTTGAGACATTTGTCTAC	1290
OY	721	ATTACCGAAATCGAACCCATCTGTGCTCTCAGAGCCTGCATGGATGGTTACAGGGGCTGA	780
Db	1291	ATTACCGAAATCGAACCCATCTGTGCTCTCAGAGCCTGCATGGATGGTTACAGGGGCTGA	1350
OY	781	AAGCTAAATGAAGTCAATCCGGCAAGTCGATGTCTATAATCTTGACAGAGAAATAGAAT	840
Db	1351	AAGCTAAATGAAGTCAATCCGGCAAGTCGATGTCTATAATCTTGACAGAGAAATAGAAT	1410
OY	841	GTAATGACACGGGAGCACTTGGATTCGATATAAAACAAGTTGTATCTATGCAATATGGGC	900
Db	1411	GTAATGACACGGGAGCACTTGGATTCGATATAAAACAAGTTGTATCTATGCAATATGGGC	1470
OY	901	CACATCAACACAGAAATCGATGTGTACACAGCGCTCCGCACTCCGGAGGTGACGTGGAGAGA	960
Db	1471	CACATCAACACAGAAATCGATGTGTACACAGCGCTCCGCACTCCGGAGGTGACGTGGAGAGA	1530
OY	961	GTACGGTCTCAGGTGACATGTCTATCTGTGGCCAGATGTGCACAAAGAGTTTCTCCTGTGCA	1020

DB 1531 GTACGTTCTAGAGTGAGCCATGTCATCTGGCCAGATGGCAACAGATGTCCTCTGCA 1590
 QY 1021 GAGGTCGTCTACTCAATTTAGCTGCTCCACAGTCCACCTTTGTTCTGTCATGACA 1080
 DB 1591 GAGGTCGTCTACTCAATTTAGCTGCTCCACAGTCCACCTTTGTTCTGTCATGACA 1650
 QY 1081 GCCACACACAGGCTTGGCAGCTGATGAACTCTATATGACCCGAGGGCGATACAG 1140
 DB 1651 GCCACACACAGGCTTGGCAGCTGATGAACTCTATATGACCCGAGGGCGATACAG 1710
 QY 1141 CAGGATGTACTGTTGCTCTAAGAAATGATGATATGATGCTGACCTGCTGCACTGCA 1200
 DB 1711 CAGGATGTACTGTTGCTCTAAGAAATGATGATGATGCTGACCTGCTGCACTGCA 1770
 QY 1201 TCATTGATGCTCCACCTTACAGAGCTGACAGATGACCAAGCAATAATCTGGACTCAAC 1260
 DB 1771 TCATTGATGCTCCACCTTACAGAGCTGACAGATGACCAAGCAATAATCTGGACTCAAC 1830
 QY 1261 AAAAATGGCCATTCAACCTAATTTATACAGATAC 1296
 DB 1831 AAAAATGGCCATTCAACCTAATTTATACAGATAC 1866

RESULT 6
 AF315687 2677 bp mRNA linear PRI:22-MAR-2002
 LOCUS Homo sapiens S-adenosylhomocysteine hydrolase-like protein mRNA,
 complete cds.
 ACCESSION AF315687
 VERSION AF315687.1 GI:16588686
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2677)
 AUTHORS Dekker, D.N., Budhla, S., Angel, N.Z., Cooper, B.J., Clark, G.J.,
 Hart, D.N. and Kato, M.
 TITLE Identification of an S-adenosylhomocysteine hydrolase-like
 transcript induced during dendritic cell differentiation
 JOURNAL Immunohematol. 53 (12): 993-1001 (2002)
 MEDLINE 21901265
 PUBMED 11904675

REFERENCE 2 (bases 1 to 2677)
 AUTHORS Dekker, D.N., Budhla, S., Angel, N.Z., Cooper, B.J., Clark, G.J.,
 Hart, D.N.J. and Kato, M.
 TITLE Direct Submission
 JOURNAL Submitted (23-OCT-2000) Dendritic Cell Research, Mater Medical
 Research Institute, Level 3, Auditory Place, South Brisbane,
 Queensland 4101, Australia

FEATURES
 source
 1..2677
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1p21"
 /clone="211(1)B"
 /cell_line="Hodgkin's disease cell line L428"
 1..114
 /note="derived from 5' RACE"
 369..1961
 /note="DCAL: dendritic cell expressed ABCY-like protein;
 expressed in dendritic cells in blood, Langerhans cells
 and dermal dendritic cells; mRNA increased during monocyte
 differentiation to monocyte-derived dendritic cells in the
 presence of GM-CSF and IL-4"
 /codon_start=1
 /product="S-adenosylhomocysteine hydrolase-like protein"
 /protein_id="AA26869.1"
 /db_xref="GI:16588687"
 /translation="MSRPDMPPLPGVEELKQAKETDAEKISPMATVTKPKQIOP
 ADMDEPTKPTLGRSLRSISQSDTSYSSAASYSDDSDSPREKQNTSKGS

SNFCVNIKIOEERREIEIAEDMSALISLRKROGERPLAGKIVGCTHTPAOTAV
 LIEETICALGACBMSACNITYSPONEEALAEVAVFAKSGSEDFEMCICDRVNM
 DGMQANNITLEDGSDILHWTYKRIPNVFKRTIRGVESYGVHLLQISRAKGCYVAM
 NVDSVYKQKFDMLYCCRESILDLKARTKIDVNGGVQVQVGVGEGCCALALAG
 AIYVITEIDICALQCMDFRVLKNEVROVDVITCGNNVYTRILDKKNSC
 IYVCMGNSIEIVTSLRPELMEERBSQVDVIMPDGRVLLLEGRLNLSCTV
 PFTVLSITATTOALALIELYNAPEDGRVKQVYLLPKMDYVSLHLPSPDHLTELT
 DDQAKYIGLUNKNPFKNRYRY"
 BASE COUNT 657 a 648 c 708 g 664 t
 ORIGIN
 Query Match 100.0%; Score 1296; DB 9; Length 2677;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACAGGCGACCAATTTCTGTGTGAAGAATCAACAGGCGAATTTGGACCGCG 60
 DB 663 TCACAGGCGACCAATTTCTGTGTGAAGAATCAACAGGCGAATTTGGACCGCG 722
 QY 61 GAGATTGAGATTGACAGACATGCTGCTGATTGACACAGAAAGCGCTAG 120
 DB 723 GAGATTGAGATTGACAGACATGCTGCTGATTGACACAGAAAGCGCTAG 782
 QY 121 GGGGAGAGCCCTTGCTGCTGCTAAATAGTGGCTGTACACACATGACAGCCAGACA 180
 DB 783 GGGGAGAGCCCTTGCTGCTGCTAAATAGTGGCTGTACACACATGACAGCCAGACA 842
 QY 181 GCGGTGTTGATTGAGACACCTGCTGCTGCGGGGCTCAGTCCGCTGCTGTTGAC 240
 DB 843 GCGGTGTTGATTGAGACACCTGCTGCTGCGGGGCTCAGTCCGCTGCTGTTGAC 902
 QY 241 ATCTACTCACTCAGAAATGAAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 DB 903 ATCTACTCACTCAGAAATGAAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962
 QY 301 GCTTGAAGGGCGAGTGAAGATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 DB 963 GCTTGAAGGGCGAGTGAAGATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
 QY 361 GATGGGTGGCAGGCCACATGATCTGATGATGATGGGGAGACCTTACCACCTGGGTTAT 420
 DB 1023 GATGGGTGGCAGGCCACATGATCTGATGATGATGGGGAGACCTTACCACCTGGGTTAT 1082
 QY 421 AAGAAGTATCCAAACGTTTAAAGATCCGAGCATTTGGAAGAAAGCGTGCTGCT 480
 DB 1083 AAGAAGTATCCAAACGTTTAAAGATCCGAGCATTTGGAAGAAAGCGTGCTGCT 1142
 QY 481 GTTCACAGGCTGATCAGCTGCCAAAGCTGGAGAGCTGCTGCTGCCCATGAAGTC 540
 DB 1143 GTTCACAGGCTGATCAGCTGCCAAAGCTGGAGAGCTGCTGCTGCCCATGAAGTC 1202
 QY 541 AATGATCTGTTACCAACAGAAATTTGATTAATCTGCTGCTGCCAGAAATCCATTG 600
 DB 1203 AATGATCTGTTACCAACAGAAATTTGATTAATCTGCTGCTGCCAGAAATCCATTG 1262
 QY 601 GATGGCTGAAGAGGACCAAGATGATGTTGGTGGAAACAAGTGGTGTGCTGCTG 660
 DB 1263 GATGGCTGAAGAGGACCAAGATGATGTTGGTGGAAACAAGTGGTGTGCTGCTG 1322
 QY 661 TATGGTGAAGTGAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB 1323 TATGGTGAAGTGAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1382
 QY 721 ATTACCGAAATGAGACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 DB 1383 ATTACCGAAATGAGACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1442
 QY 781 AAGCTAAATGAAGTCATCCGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 1443 AAGCTAAATGAAGTCATCCGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1502
 QY 841 GTAGGACAGGGAGACCTTGTGATGCTGATAAACAATTTGATGCTGCTGCTGCTGCTGCTGCT 900

Db 1503 GTAGTACACGCGGAGCACTTGGATCGCATGAAAAACAGTTGATCGTATGCAATATGGGC 1562
 Qy 901 CACTCCAAACAGAAATCGATGTAGCAGCCTCCGACCTCCGAGCTGACGTGGAGCGA 960
 Db 1563 CAGTCCAAACAGAAATCGATGTAGCAGCCTCCGACCTCCGAGCTGACGTGGAGCGA 1622
 Qy 961 GTACGTTCTCAGGTGGACCATGTCTATCTGGCCAGATGGCAAAAGATTGCTCTCTGGCA 1020
 Db 1623 GTACGTTCTCAGGTGGACCATGTCTATCTGGCCAGATGGCAAAAGATTGCTCTCTGGCA 1682
 Qy 1021 GAGGCTGTCTACGATTTAGTGGCTCCAGAGTCCACCTTTGTCCTGTCATGCA 1080
 Db 1683 GAGGCTGTCTACGATTTAGTGGCTCCAGAGTCCACCTTTGTCCTGTCATGCA 1742
 Qy 1081 GCCACACACAGCGCTTGGCAGCTGATAGACTGATATGACCCGAGGCGGATGACAG 1140
 Db 1743 GCCACACACAGCGCTTGGCAGCTGATAGACTGATATGACCCGAGGCGGATGACAG 1802
 Qy 1141 CAGATGTGTACTTGTCTTCCATAGAAAATGATGATACGTTGGCAGCTTGCATCTGCCA 1200
 Db 1803 CAGATGTGTACTTGTCTTCCATAGAAAATGATGATACGTTGGCAGCTTGCATCTGCCA 1862
 Qy 1201 TCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGAAAATATCTGGAGCTCAAC 1260
 Db 1863 TCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGAAAATATCTGGAGCTCAAC 1922
 Qy 1261 AAAAATGGCCATTCACAACTTAATTATTACAGATAC 1296
 Db 1923 AAAAATGGCCATTCACAACTTAATTATTACAGATAC 1958

RESULT 7
 HSM800298
 LOCUS
 DEFINITION Homo sapiens mRNA; cDNA DKFZp564A1523 (from clone DKFZp564A1523); partial cds.

ACCESSION AL049954
 VERSION AL049954.1 GI:4884203
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.

REFERENCE
 AUTHORS Wambuit, R., Heuener, D., Nemes, H.W., Gassenhuber, J. and Wiemann, S.
 TITLE Direct Submision
 JOURNAL Submitted (15-MAY-1999) MIPS, Am Klopferstr. 18a, D-82152 Martinsried, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Emmeil s. Wiemannedkz-helidelberg.de; sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
 This clone (DKFZp564A1523) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://www.mips.biochem.mpg.de/proj/cDNA/.

FEATURES
 source
 1..2510
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="1, 15.55 cR from DIS248"
 /clone="DKFZp564A1523"
 /tissue="brain"
 /clone_id="564 (synonym: hfbz2). Vector pAMP1; host XL1-blue; sites NotI + SalI"
 /dev_stage="fetal"
 1..1796
 /gene="DKFZp564A1523"
 /note="S-adenosyl homocysteine hydrolase homolog XPVKona splice variant"

gene
 CDS
 1..1796
 /gene="DKFZp564A1523"
 /note="S-adenosyl homocysteine hydrolase homolog XPVKona splice variant"

/codon_start=3
 /product="hypothetical protein"
 /protein_id="CA43223.1"
 /db_xref="GI:4884204"
 /translation="GILTVELSSQVCVYILLILHRHSGSRNGKICAEISOPALPLPOE
 TOAPSGSENLRLNRGRRRAESSONLSLRVYACNGNSRDMRWITOLIMPEIDSL
 ISCFQIQQFNDDMOETKFFPTKRRRLSLISQSSSTDSISAASTYDSDDEVSPE
 KOQNSKNSFNQCVKNIKQAEFRRELEISQDSALSLKRRQGEKPLAGATIVC
 THITACAVALEIETLCALGAOC RMSGCNLYSTONVAAALAEVAVAMNGESDDFW
 WCIDRCVNMGOAMNMLIDGGLTHWYKKYPMVFKIRIGIVSEYGVARYQLSK
 AGKLCVPMNVNDSEVKOKFDNLYCRESLIDGLRTDVAFGGKQVYGVGKVG
 CCAATKALGALVYETEDPTICALQACMDGPPVYVNLNIVRDVYVICTGKNVYVE
 HDBRMKSCIVCMNGHSENTEIDVTSLEPTETRYRVSQDHYVMPGKRVYLLAEGR
 LNLNLSCTVPEPVLSTATVQALALIEVYNPEGRYKQDYVLLPKRDEYASLHLPS
 FDLNLTETLDOAYVILGLNNKPPKPYRR"
 polyA_site 2490
 BASE COUNT 664 a 556 c 618 g 672 t
 ORIGIN
 Query Match 99.9%; Score 1294.4; DB 9; Length 2510;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1295; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 TCCAAGGCGAGCAGCAATTTCTGTGTGAAGAACATCAAGCAGAGCAATTTGAGCGCG 60
 Db 498 TCCAAGGCGAGCAGCAATTTCTGTGTGAAGAACATCAAGCAGAGCAATTTGAGCGCG 557
 Qy 61 GAGATTGAGATTGACAGCAAGACATGCTGCTGATTTCACTCAGAAACGTGCTGAG 120
 Db 558 GAGATTGAGATTGACAGCAAGACATGCTGCTGATTTCACTCAGAAACGTGCTGAG 617
 Qy 121 GGGGGAAGCCCTTGGCTGGGTCTAAATATGAGGCTGTACACACATCAAGCCAGACA 180
 Db 618 GGGGGAAGCCCTTGGCTGGGTCTAAATATGAGGCTGTACACACATCAAGCCAGACA 677
 Qy 181 GCGGTGTTGATTGACACACTGTGCTGGGGGCTCAGTGGCGCTGCTCTGTAAAC 240
 Db 678 GCGGTGTTGATTGACACACTGTGCTGGGGGCTCAGTGGCGCTGCTCTGTAAAC 737
 Qy 241 ATCTACTCACTCAGAAATGAAATGAACTGACAGCTGAGCTGAGGTGAGTGTTC 300
 Db 738 ATCTACTCACTCAGAAATGAAATGAACTGACAGCTGAGCTGAGGTGAGTGTTC 797
 Qy 301 GCTTGAAGGCGGATGACAGAGATGATCTGTGTGTGTATTGACCGCTGTGTAAACATG 360
 Db 798 GCTTGAAGGCGGATGACAGAGATGATCTGTGTGTGTATTGACCGCTGTGTAAACATG 857
 Qy 361 GATGGGTGCGAGGCGCAATGATCTGTGATGATGGGAGACTTAACCACTGGGTTAT 420
 Db 858 GATGGGTGCGAGGCGCAATGATCTGTGATGATGGGAGACTTAACCACTGGGTTAT 917
 Qy 421 AAGAAATATCCAAACGTGTTAAGAAATCGAGCAATTTGGAGAGAGCGTACTGT 480
 Db 918 AAGAAATATCCAAACGTGTTAAGAAATCGAGCAATTTGGAGAGAGCGTACTGT 977
 Qy 481 GTTCACAGCGTATCACCTCTCCAAAGCTGGGAAGCTGTGTCGGGCCATGAGACGC 540
 Db 978 GTTCACAGCGTATCACCTCTCCAAAGCTGGGAAGCTGTGTCGGGCCATGAGACGC 1037
 Qy 541 AATGATTTGTTACCAACAGAAATTTGATTAATCTGTACTGCTGCCAGAAATCCATTTTG 600
 Db 1038 AATGATTTGTTACCAACAGAAATTTGATTAATCTGTACTGCTGCCAGAAATCCATTTTG 1097
 Qy 601 GATGGCTGAAGAGGACACACAGATGATGTTGGTGGAAACAGTGGTGTGTGGC 660
 Db 1098 GATGGCTGAAGAGGACACACAGATGATGTTGGTGGAAACAGTGGTGTGTGGC 1157
 Qy 661 TATGTTGAGTGAAGCAAGGCGTGTGCTGCTCAAAAGCTCTTGAGCAATTTGTAC 720
 Db 1158 TATGTTGAGTGAAGCAAGGCGTGTGCTGCTCAAAAGCTCTTGAGCAATTTGTAC 1217
 Qy 721 ATTACGAATGACCCCATGTGCTGTCAGGCGCTGATGATGGATTGAGGTGATG 780

Db 1218 ATTACGAATTCAGCCCATCTGTCTGTGAGGCTGATGATGAGGTTCAGGGTGTGTA 1277
 QY 781 AAGCTAAATGATGATCCGGCAAGTCGATGCTGTAATTAATTCACAGAAATPAGAT 840
 Db 1278 AAGCTAAATGATGATCCGGCAAGTCGATGCTGTAATTAATTCACAGAAATPAGAT 1337
 QY 841 GTATGACAGGAGGACCTTGATGATGATGATGATGATGATGATGATGATGATGAT 900
 Db 1338 GTATGACAGGAGGACCTTGATGATGATGATGATGATGATGATGATGATGATGAT 1397
 QY 901 CACCCAAACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 Db 1398 CACTCCAAACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457
 QY 961 GTACGTTCTCAGTGAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 Db 1458 GTACGTTCTCAGTGAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1517
 QY 1021 GAGGCTGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 Db 1518 GAGGCTGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1577
 QY 1081 GCCCAACAGAGGCTTGGGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 Db 1578 GCCCAACAGAGGCTTGGGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1637
 QY 1141 CAGGATGCTACTGCTTCTCTAAGAAATGATGATGATGATGATGATGATGATGATGAT 1200
 Db 1638 CAGGATGCTACTGCTTCTCTAAGAAATGATGATGATGATGATGATGATGATGATGAT 1697
 QY 1201 TCATTGATGCTCACTTCAAGAGCTGACAGATGACAGCAAGCAAAATATCTGGACTCAAC 1260
 Db 1698 TCATTGATGCTCACTTCAAGAGCTGACAGATGACAGCAAGCAAAATATCTGGACTCAAC 1757
 QY 1261 AAAAATGGCCATCAAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1296
 Db 1758 AAAAATGGCCATCAAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1793
 RESULT 8
 BC018218
 LOCUS 1772 bp mRNA linear ROD 07-AUG-2002
 DEFINITION Mus musculus, S-adenosylhomocysteine hydrolase-like 1, clone
 BC018218
 ACCSSION MGC:18748 IMAGE:4007102, mRNA, complete cds.
 VERSION BC018218.1 GI:17390492
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1772)
 Strausberg, R.
 Direct Submission
 Submitted (03-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-rt@mail.nih.gov
 Reissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M.,
 Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRK Plate: 23 Row: n Column: 6
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Genomescan gene prediction, Similarity but not identity
 to protein.

FEATURES

source

CDS

1..1772
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="C2E0H II"
 /clone="MGC:18748 IMAGE:4007102"
 /tissue_type="Mammary tumor metastasized to lung."
 /MVT-LTR/Mnt1 model. Expression driven by an MVT-LTR
 enhancer."
 /clone_id="NCI CGAP Lu30"
 /lab_host="DH10B"
 /note="vector: pcMV-Sport6"
 191..1693
 /codon_start=1
 /product="S-adenosylhomocysteine hydrolase-like 1"
 /protein_id="AAH18218.1"
 /db_xref="GI:17390493"
 /db_xref="locus:10768"
 /translation="MAYVYKAPKQIDPADDMQETFPPTKTRRSLSRSISQSTDS
 YSSAASYTDSDDVSPREKQTSKSSNECVNINIOAERGRELEIAEDMALLS
 LKRAQGEKPLAGAKIVGCTHTIQTAVLLETLALAOCSMSACNITSEVEMAAAL
 AEAVAVPAMGSEDEDFMNCIDRCVNMNDQAMMIDDDGLTHMYVKKYPPFKKI
 RGYEESVTVGYVHRYLSKSKGLRCPVNMVNDVSTKQKFDNLVYCCRSIIDLKRTTD
 VMFGKQVYVCGYREVGKCCALKALGATYITETIDPICALQACDGFERYKNEYI
 RQDVVITCTGKNVAVREHLDNRKNSCTVCMNSVEIDVLSLRPELWERSQ
 RDVHVIWPGKRVVILABGRILNLSCTVPEVLSITATQALALIEYLNMEGRYKOD
 VYLLPKMDEVASLHLPSFDAHYTELDQAKYLGLINKNPFKPNRYR"

BASE COUNT 452 a 406 c 506 g 398 t
 ORIGIN

Query Match 90.0%; Score 1166.4; DB 10; Length 1772;
 Best Local Similarity 93.8%; Pred. No. 0;
 Matches 1215; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 1 TCAGAGGCGAGACGAATTTCTGTGTGAAGACATCAAGCAGGCAATTTGAGCGCGG 60
 Db 395 TCAGAGGCGAGACGAATTTCTGTGTGAAGACATCAAGCAGGCAATTTGAGCGCGG 454
 QY 61 GAGATTGAGATTGACAGACATGCTGCTGATTTCACTCAGGAAACGCTCTAG 120
 Db 455 GAGATTGAGATTGACAGACATGCTGCTGATTTCACTCAGGAAACGCTCTAG 514
 QY 121 GGGGAGAAAGCCCTTGCTGCTGCTAAATAGTGGCTGTACACATCAGACCCAGACA 180
 Db 515 GGGGAGAAAGCCCTTGCTGCTGCTAAATAGTGGCTGTACACATCAGACCCAGACA 574
 QY 181 GCGGTGTTGATTGAACACATCTGTGCTGGGGGCTCAGTGGCTGCTGCTGTATAC 240
 Db 575 GCGGTGTTGATTGAACACATCTGTGCTGGGGGCTCAGTGGCTGCTGCTGTATAC 634
 QY 241 ATCTACCACTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 Db 635 ATCTACCACTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 694
 QY 301 GCTTGAAGAGGAGTACAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 Db 695 GCTTGAAGAGGAGTACAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 754
 QY 361 GATGGGTGGCAGGCAACATGATCTGATGATGATGATGATGATGATGATGATGATGAT 420
 Db 755 GATGGGTGGCAGGCAACATGATCTGATGATGATGATGATGATGATGATGATGATGAT 814
 QY 421 AAGAGTATCCAAAGGTGTTAAGAAATCAGAGGATGTTGGAAGAGAGAGCTGATGAT 480
 Db 815 AAGAGTATCCAAAGGTGTTAAGAAATCAGAGGATGTTGGAAGAGAGAGCTGATGAT 874


```

OY 301 GCTTGGAAAGGCGAGTCAAGATGACTTCTGGTGTATGTAACCGCTGTGTGAACATG 360
Db 862 GCGTGGAGGAGAGATGACATGTTGGTGTATGATGATGATGATGATGATGATGATG 921
OY 361 GATGGGTGGAGGCGCAATATGATCTGATGATGATGATGATGATGATGATGATGATG 420
Db 922 GAGGGGTGGAGGCGCAATATGATCTGATGATGATGATGATGATGATGATGATGATG 981
OY 421 AAGAAGTATCCAAACGTGTTTAAAGATGATGATGATGATGATGATGATGATGATGAT 480
Db 982 AAAAAGTATCCAAACGTGTTTAAAGATGATGATGATGATGATGATGATGATGATGAT 1041
OY 481 GTTCAAGGCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 1042 GTTCAAGGCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1101
OY 541 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 1102 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
OY 601 GATGGCTGGAAGAGACCAAGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 1162 GATGGCTGGAAGAGACCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1221
OY 661 TATGGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 1222 TATGGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1281
OY 721 AATGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 1282 AATGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1341
OY 781 AAGCTAAATGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 1342 AAGCTAAATGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1401
OY 841 GTATGACACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 1402 GTATGACACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1461
OY 901 CACTCAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 1462 CACTCAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1521
OY 961 GTACGTTCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 1522 GTACGTTCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1581
OY 1021 GAGGCTGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1582 GAGGCTGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1641
OY 1081 GCCACAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1642 GCTTCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1701
OY 1141 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1702 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1761
OY 1201 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1762 ACCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1821
OY 1261 AAAAATGGGCGCATCAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1295
Db 1822 AAGAATGGGCGCATCAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1856

```

```

RESULT 12
BC024325
LOCUS BC024325 5052 bp MRNA linear PRI 12-MAR-2002

```

```

DEFINITION Homo sapiens, KIA0828 protein, clone MGC:21525 IMAGE:3907552,
mRNA, complete cds.
ACCESSION BC024325
VERSION BC024325.1 GI:19353091
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 5052)
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mhc.nci.nih.gov
Contact: MGC help desk
Email: gcgbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mdcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Gilwood, J., Rodriguez, A., and Myers,
R. M.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAC Plate: 22 Row: 1 Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: similarity but not identity to protein.

FEATURES

source

1..5052

location/Qualifiers

1..5052

organism="Homo sapiens"

db_xref="locusID:23382"

db_xref="taxon:9606"

clone="MGC:21525 IMAGE:3907552"

tissue="uterus"

clone_lib="NIH_MGC_71"

lab_host="DH10B"

note="Vector: pCMV-SPORT6"

55..1890

codon_start=1

product="KIA0828 protein"

protein_id="AAH24325.1"

db_xref="GI:19353092"

translation="MSYQVYVSAALAAKPYEELKDLSPSEASQGLSTAAGMAAP
AGGDPPEAPAAERPPVPGSGSPAALSPAAGVPAASMAKRSDDPHORHDDG
EALVSPDSTEARPRYKQIOPADQKQEFKRPKIGRSLSSISQSDSSAA
SYTSSDDETPRDKQKNSGSSDFCKNIKQAEGRRELEIAEOPALMALAKRA
QGERPLGAKIVGCHITQAIVLMTGALACQRMALACNISTLNVAALAESGR
PVPAMKGSDEDFWQICDRVNAVNDVTKQFNDLCCRESILYDKRTPMKRIGIVE
ESYGVHRLYQLSRAKLCVPAVMNDVTKQFNDLCCRESILYDKRTPMKRIGIVE
KQYVVGGEVTEGCAALKAMGSLVYETEDPCALCAQSGRLVNLVITQDVI
VITQGNKNTVTRHRLDMKRSCTVNCNNGVSTATTATDALILILNAPSRVYKQVYILP
WPKSRVILLAEGLINLSCSTVPEFVISTATTATDALILILNAPSRVYKQVYILP
KKMDEYVASLTPFDALTELTEDQALYLGNNGPKPYRYR"

BASE COUNT 1300 a 1184 c 1145 g 1423 t

ORIGIN

Query Match 66.5% Score 861.4; DB 9; Length 5052;
Best Local Similarity 79.1% P-Val 1.4e-257;
Matches 1024; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

```

OY 1 TCCAAGGCGAGCAATTTCTGTGTAAGATCAATCAAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 592 TCTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 651
OY 61 GAGATTGAGATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120

```


Db	Accession	Gene	Species	Length
Db	824	AAGAAATACCCCGACTACTTCAAGGCGCATCGGGGCGATGCTGGAGGAAGTGTGACCGGG		883
Qy	481	GTTCAAGGCTGTATACGCTCTCCAAAGCTGGGAAGCTGTGTTCGGCCATGAACGTC		540
Db	884	GTCACACCGGCTGTACATGCTGTCAAAAGGCGGAAAACTTACTGTTCGGCCATCAACGTT		943
Qy	541	AATGATTCGTACCAAAAGAAAGTTGTAATCTGTACTGCTGCCGAATCCATTG		600
Db	944	AACGACTACGTACCAAGAAAGTTGTAATCTGTACTGCTGCCGAATCCATTG		1003
Qy	601	GATGCGCTGAAGAGACCAAGATGTGATGTTGGTGGGAAGAAAGTGTGTGTGTC		660
Db	1004	GACATGTGAAAGGACCAACGAGATATTAATGTTGGGGGAAGACAGTGTGTGTGTC		1063
Qy	661	TATGCTGAAGTATGCGAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT		720
Db	1064	TACGCTGTATGTGGGAAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT		1123
Qy	721	ATTACCGAAATGAGACCCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT		780
Db	1124	GTTACGGAATGTGATCCCATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT		1183
Qy	781	AAGCTAAATGAAGTATCGCGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT		840
Db	1184	CGGCTCAAGAGTATCATGAGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT		1243
Qy	841	GTAATGACACGAGGACCTTGATGCTGATGAAGAAAGTGTGTGTGTGTGTGTGT		900
Db	1244	GTTATACAGGAGATCATGATGATGATGAAGAAAGTGTGTGTGTGTGTGTGTGT		1303
Qy	901	CACCTCAACAGAAATGATGTGACAGCGCCCGCATCTCGGAGGTGACGTGGAGCG		960
Db	1304	CATTCTGT		1363
Qy	961	GTAATGACAGTGTGAGCATGATGATGATGATGATGATGATGATGATGATGATG		1020
Db	1364	GTCGCTGT		1423
Qy	1021	GAGGCTGT		1080
Db	1424	GAGGGAAGAGT		1483
Qy	1081	GCCAAACAGAGGCTTTGGCACTGATTAACATCTTAATGACACCGGAGGCGATTA		1140
Db	1484	TCATCCACCCAGGCTTTGGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT		1540
Qy	1141	CAGATGT		1200
Db	1541	TGCGATGT		1600
Qy	1201	TCATTTGATGCGCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAC		1260
Db	1601	ACCTTCGATGCGCATCTCACGAGAGCTCACAGATGACAGTCAAGTTATGSGCTTA		1660
Qy	1261	AAAAATGGGCAATTCAAACTTAATATTATACAGATA		1295
Db	1661	AAGCGCGGCGCTTTAAAGCAATTACTACAGATA		1695
RESULT 15	BC003631			
LOCUS	BC003631			
DEFINITION	Homo sapiens, Similar to S-adenosylhomocysteine hydrolase-like 1,			
ACCESSION	BC003631			
VERSION	BC003631.1			
KEYWORDS	GI:13177702			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 1323)			

FEATURES	source
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT	Contact: MGC help desk Email: cgapbs-rt@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbio.org contact: amadansystemsbiology.org Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettelman and Anuradha Madan
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 3 Row: 6 Column: 6. Location/Qualifiers 1. 1323 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3010755" /tissue_type="Muscle, rhabdomyosarcoma" /clone_id="NIH_MGC_17" /lab_host="DH10B-R" /note="Vector: pOTB7" <1. 611 /codon_start=3 /product="Similar to S-adenosylhomocysteine hydrolase-like 1" /protein_id="AAH03631.1" /db_xref="GI:13177703" /translation="HEBALAIYITRIPDICALQCMDSFRRYKLNERYROYDVYITCTGRVNVREHLDKRNKNSCTIVCMGMSNTEIDVTSRTBELTWERVRSDVHIVPDGKRVLVLAERLNLSCSTVPTFVLSTATQTLALIELYNAEGRYKODVYLLPKRMDEVYASIHLPSPFAHLTEITDDQAKYGLNKNGPFRKPYRRY"
BASE COUNT	365 a 303 c 271 g 384 t
ORIGIN	
Query Match	46.2%; Score 599; DB 9; Length 1323;
Best Local Similarity	100.0%; Pred. No. 1.4e-175;
Matches 599; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY 698	AAGCTCTTGGAGCAATGTCTACATTTACCGAAATCGACCCCATCTGTGCTTCGAGGCT 757
Db 10	AAGCTCTTGGAGCAATGTCTACATTTACCGAAATCGACCCCATCTGTGCTTCGAGGCT 69
QY 758	GCATGATGGATTGAGGTTGTTAAAGCTAAATGAAGTCATCCGGCAAGTCATGCTGTA 817
Db 70	GCATGATGGATTGAGGTTGTTAAAGCTAAATGAAGTCATCCGGCAAGTCATGCTGTA 129
QY 818	TAACTTGCACAGAAATTAAGATGTAGTGACACGGGAGCACTTGGATCGATGAAAAACA 877
Db 130	TAACTTGCACAGAAATTAAGATGTAGTGACACGGGAGCACTTGGATCGATGAAAAACA 189
QY 878	GTTGATCGTATGCAATATGAGGCGACATCCACAGAAATCGATGTGACAGCCTCCGCA 937
Db 190	GTTGATCGTATGCAATATGAGGCGACATCCACAGAAATCGATGTGACAGCCTCCGCA 249
QY 938	CTCCGGAGCTGACCTGGAGAGGATAGGTTTTCACGGTGGACATGTATCTGGCAATG 997
Db 250	CTCCGGAGCTGACCTGGAGAGGATAGGTTTTCACGGTGGACATGTATCTGGCAATG 309
QY 998	GCAACGAGTGTCTCTCTGGAGAGGCTCTTACCAATTTAGCTGGCTCCACAGTTG 1057
Db 310	GCAACGAGTGTCTCTCTGGAGAGGCTCTTACCAATTTAGCTGGCTCCACAGTTG 369
QY 1058	CCACCTTTGTTCTGTCATCACGCCACACACAGAGGCTTGGCACTATAGAACTCTATA 1117

Db 370 CCACCTTGTCTGTCATCACAAGCCACAACAGGCTTGGCAGTGAATACTCTATA 429
QY 1118 ATGCACCCGAGGGGCGATACAGAGAGATGTACTTGTCTCTAAGAAAATGATGAAT 1177
|||||
Db 430 ATGCACCCGAGGGGCGATACAGAGAGATGTACTTGTCTCTAAGAAAATGATGAAT 489
QY 1178 ACGTTCAGCTGCATCTGCCATCTTGAATGACCCACCTTACAGAGCTGACAGATGACC 1237
|||||
Db 490 ACGTTCAGCTGCATCTGCCATCTTGAATGACCCACCTTACAGAGCTGACAGATGACC 549
QY 1238 AAGCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAACCTAATTATTACAGATAC 1296
|||||
Db 550 AAGCAAAATATCTGGGACTCAACAAAATGGGCCATTCAAACCTAATTATTACAGATAC 608

Search completed: April 21, 2003, 02:39:26
Job time : 3367.48 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 21:15:39 ; Search time 275.45 Seconds
(without alignments)
10595.712 Million cell updates/sec

Title: US-09-782-051-1_COPY_549_1844

Perfect score: 1296
Sequence: 1 tccaagggcagcagcaattt.....aaccaattatcacagatcac 1296

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq.101002.*

1:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
23:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
24:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1296	100.0	2563	19	AAV28617
2	1296	100.0	3634	23	AAV23195
3	1296	100.0	3634	23	AAV29032
4	1294.4	99.9	3636	20	AAV73924
5	815.8	62.9	3030	23	AAV76216
6	676.6	52.2	1566	23	ABL03693
7	628.4	48.5	1776	23	ABL29787
8	486.2	37.5	4406	23	ABL03692
9	461.4	35.6	553	24	ABL79045

C	10	422	32.6	504	23	ABV44141	Human prostate exp
C	11	406.6	31.4	5537	23	ABT29786	Drosophila melanog
C	12	406.4	31.0	6636	21	AAV80766	Human colon cancer
	13	402.4	31.0	2200	21	AAV7823	Human cancer assoc
	14	402.4	31.0	2658	23	ABV23126	Human prostate exp
	15	402.4	31.0	2658	23	ABV28967	Human prostate exp
	16	378.2	29.2	1592	23	ABT28157	Drosophila melanog
	17	363.8	28.1	435	23	ABV14215	Human prostate exp
	18	337	26.0	384	24	ABV93848	Gene #346 used to
	19	314.4	24.3	2211	22	AAV28156	Human S-adenosylho
	20	313.2	24.2	4380	23	ABT28156	Drosophila melanog
	21	245	18.9	939	23	AAV76217	DNA encoding novel
	22	244.8	18.9	312	23	ABV35309	Human prostate exp
	23	242.8	18.7	1727	21	AAV46315	Arabidopsis thalia
	24	242.6	18.7	278	24	ABV73754	Bovine embryonic g
	25	241.6	18.6	1461	22	AAV09278	Trichomonas vagina
	26	236.2	18.2	1488	21	AAV39567	M. tuberculosis an
C	27	236.2	18.2	4411529	22	AAV19682	Mycobacterium tube
	28	233.6	18.0	1461	22	AAV09277	Trichomonas vagina
	29	231.8	17.9	1728	21	AAV39886	Arabidopsis thalia
	30	228.8	17.7	1865	21	AAV44037	Zea mays DNA fragm
	31	226.4	17.5	1494	21	AAV46817	Arabidopsis thalia
	32	225.6	17.4	1422	22	AAV65801	C glutamylum codin
	33	225.6	17.4	1557	22	AAV71862	Corynebacterium gl
	34	225.6	17.4	1557	23	AAV6121	C. glutamylum gene
	35	225.6	17.4	1939	24	ABV52894	Sah gene encoding
	36	225.6	17.4	349980	22	AAV68526	C glutamylum codin
	37	221	17.1	1812	17	AAV29194	S-adenosylhomocyst
	38	217.8	16.8	1812	13	AAV29194	Nicotiana tabacum
	39	214.6	16.6	1767	17	AAV44513	Aspergillus oryzae
	40	205.8	15.9	58909	22	AAV28543	Genomic fragment #
C	41	202	15.6	656	21	AAV14585	Aspergillus oryzae
	42	190	14.7	1877	18	AAV06354	Al-2-Long (Al-21)
C	43	189.2	14.6	1941	17	AAV44515	Arabidopsis S-aden
	44	186	14.4	1396	22	AAV1864	Corynebacterium gl
	45	186	14.4	1396	23	AAV5123	C. glutamylum gene

ALIGNMENTS

RESULT 1	AAV28617	standard; CDNA; 2563 BP.
ID	AAV28617	
XX	AAV28617	
AC	AAV28617	
XX	AAV28617	
DE	21-AUG-1998	(first entry)
XX	21-AUG-1998	
DE	Nucleotide sequence of the ODD4b5.3 enzyme.	
XX	DD4b9.3 enzyme; S-adenosyl-L-homocysteine hydrolase-type activity;	
KW	ACY-type activity; stimulation; inhibition; dendritic cell; cancer;	
KW	autoimmune disease; transplantation; ss.	
XX		
OS	Homo sapiens.	
XX		
EH	Key	Location/Qualifiers
FT	CDS	3..1847
FT		/*tag- a
FT		/product- "DD4b5.3 enzyme"
XX		
PN	W09814562-A1	
PD	09-APR-1998.	
XX		
PF	06-OCT-1997;	97NO-NR00133.
XX		
PR	04-OCT-1996;	96NZ-0299507.
XX		
PA	(HART/) HART D N J.	
XX		
PI	Hart DNT;	

XX WPI: 1998-240074/21.
 DR P-PSDB: AAM56097.
 PT New isolated S-adenosyl-L-homocysteine hydrolase enzyme - is used to
 develop products which can be used in the treatment of e.g.
 PT auto-immune disease, transplantations or cancers
 PS Claim 7, Fig 1, 33pp; English.
 CC This is the nucleotide sequence of the DD4b9.3 enzyme which has
 CC S-adenosyl-L-homocysteine hydrolase (AHcy)-type activity. Its
 CC products can be used to identify substances which have a stimulatory
 CC or inhibitory effect on the enzyme activity. Such substances can be
 CC used to modulate dendritic cell (DC) function and for immunomodulation.
 CC They can be used in the treatment of e.g. autoimmune diseases,
 CC transplantations or cancers. The products can also be used for
 CC detection and diagnosis.
 CC
 XX

Sequence: 2563 BP; 646 A; 604 C; 677 G; 636 T; 0 other;

Query Match 100.0%; Score 1296; DB 19; Length 2563;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACAGGCGACGACGATTTCTGTGTAAGAACATCAAGCGGCAATTTGGAGCCGG 60
 DB 549 TCACAGGCGACGACGATTTCTGTGTAAGAACATCAAGCGGCAATTTGGAGCCGG 608
 QY 61 GAGATTGATTCAGACAGACATGCTGCTGATTTCTGACGAGAAACGCTCAG 120
 DB 609 GAGATTGATTCAGACAGACATGCTGCTGATTTCTGACGAGAAACGCTCAG 668
 QY 121 GGGGAGAACCCCTTGGCTGTCTAAATAGTGGGCTGTACACATCAGACCCAGACA 180
 DB 669 GGGGAGAACCCCTTGGCTGTCTAAATAGTGGGCTGTACACATCAGACCCAGACA 728
 QY 181 GCGGTGTGATTGAGACACTGTGTGCTGGGGGCTCAGTGGCGTGTCTGTGTAC 240
 DB 729 GCGGTGTGATTGAGACACTGTGTGCTGGGGGCTCAGTGGCGTGTCTGTGTAC 788
 QY 241 ATCTACTACCTCAGAAATGAGATAGCTGACAGACTGGCTGAGGCTGAGTTGCTTC 300
 DB 789 ATCTACTACCTCAGAAATGAGATAGCTGACAGACTGGCTGAGGCTGAGTTGCTTC 848
 QY 301 GCTTGAAGGGGAGCTCAGAAAGATGACTTGTGTGTGTATTGACCGCTGTGTGAACATG 360
 DB 849 GCTTGAAGGGGAGCTCAGAAAGATGACTTGTGTGTGTATTGACCGCTGTGTGAACATG 908
 QY 361 GATGGGTGGCAGGCCAATGATCTGTGATGATGGGGAGACTTAACCACTGGGTTAT 420
 DB 909 GATGGGTGGCAGGCCAATGATCTGTGATGATGGGGAGACTTAACCACTGGGTTAT 968
 QY 421 AAGAAGTATCCAAACGTTTAAAGATCCGAGCATTTGGAAGAGAGCTGACTGAT 480
 DB 969 AAGAAGTATCCAAACGTTTAAAGATCCGAGCATTTGGAAGAGAGCTGACTGAT 1028
 QY 481 GTTCACAGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGGCCATGAACGTC 540
 DB 1029 GTTCACAGCTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGGCCATGAACGTC 1088
 QY 541 AATGTTCTTTTACCAACAGAGTTTATTAATCTGTACTGTCCGAGAAATCCATTTTG 600
 DB 1089 AATGTTCTTTTACCAACAGAGTTTATTAATCTGTACTGTCCGAGAAATCCATTTTG 1148
 QY 601 GATGCGCTGAAGAGGACACAGATGTGATGTTTGGTGGAAACAAGTGGTGTGTGCG 660
 DB 1149 GATGCGCTGAAGAGGACACAGATGTGATGTTTGGTGGAAACAAGTGGTGTGTGCG 1208
 QY 661 TATGCTGAGTAGGAGGAGGCTGTGTCTCTCAAAAGCTCTTGAGCAATTTGTCTAC 720
 DB 1209 TATGCTGAGTAGGAGGAGGCTGTGTCTCTCAAAAGCTCTTGAGCAATTTGTCTAC 1268

QY 721 ATTACCGAATCGACCCCATCTGTCTGTGACAGGCGCTCATGATGAGGTTCAGGGTGA 780
 DB 1269 ATTACCGAATCGACCCCATCTGTCTGTGACAGGCGCTCATGATGAGGTTCAGGGTGA 1328
 QY 781 AAGCTAATGAAGTATCATCGGCAAGTGTGATGCTGATTAATCTGACAGGAATTAAGAT 840
 DB 1329 AAGCTAATGAAGTATCATCGGCAAGTGTGATGCTGATTAATCTGACAGGAATTAAGAT 1388
 QY 841 GTAGTGACAGGGGACACTTGGATGCAATGAANAACAGTTGATGTATGCAATTTGGCC 900
 DB 1389 GTAGTGACAGGGGACACTTGGATGCAATGAANAACAGTTGATGTATGCAATTTGGCC 1448
 QY 901 CACTCAACACAGAAATGATGTGACAGCCCTCCGACATCCGGAGCTGACGTGGAGGGA 960
 DB 1449 CACTCAACACAGAAATGATGTGACAGCCCTCCGACATCCGGAGCTGACGTGGAGGGA 1508
 QY 961 GTACGTTCTCAGGTGACCATGTCATCTGCGCAGATGGCAAAAGATTGCTCCTGGCA 1020
 DB 1509 GTACGTTCTCAGGTGACCATGTCATCTGCGCAGATGGCAAAAGATTGCTCCTGGCA 1568
 QY 1021 GAGGTGCTTACTCAATTTGAGCTGTCCACAGTTCCACCTTTGTTCTGTCTACATA 1080
 DB 1569 GAGGTGCTTACTCAATTTGAGCTGTCCACAGTTCCACCTTTGTTCTGTCTACATA 1628
 QY 1081 GCCACAACAGGCTTGGCACTGATAGAACTCTTAATGCAACCGGAGGGGATACAG 1140
 DB 1629 GCCACAACAGGCTTGGCACTGATAGAACTCTTAATGCAACCGGAGGGGATACAG 1688
 QY 1141 CAGGATGTACTCTTCTTAAAGAAATGATAGTATGCTGCCAGCTTGCATCTGCA 1200
 DB 1689 CAGGATGTACTCTTCTTAAAGAAATGATAGTATGCTGCCAGCTTGCATCTGCA 1748
 QY 1201 TCATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACTCAAC 1260
 DB 1749 TCATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACTCAAC 1808
 QY 1261 AAAATGGGCACTTAAACCTAATTAATACAGATAC 1296
 DB 1809 AAAATGGGCACTTAAACCTAATTAATACAGATAC 1844

RESULT 2
 ABV23195
 ID ABV23195 standard; cDNA; 3634 BP.
 XX
 AC ABV23195;
 XX
 DE 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 23186.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-18319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-21314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI, 2001-662795/76.
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 XX Claim 1, Page 6147-6148; 11750pp; English.
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV6213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;
 Query Match 100.0%; Score 1296; DB 23; Length 3634;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCACAGGCGACAGCAATTTCTGTGAAGAATCAAGAGGAGATTTGAGCCGG 60
 DB 619 TCACAGGCGACAGCAATTTCTGTGAAGAATCAAGAGGAGATTTGAGCCGG 678
 QY 61 GAGATTGAGTTCAGAGCAACATGCTGCTGATTTCACTCAGAAACGCTCTAG 120
 DB 679 GAGATTGAGTTCAGAGCAACATGCTGCTGATTTCACTCAGAAACGCTCTAG 738
 QY 121 GGGGAGAGCCCTGGGCTGCTAAATAGTGGCTGACACATACAGCCAGCA 180
 DB 739 GGGGAGAGCCCTGGGCTGCTAAATAGTGGCTGACACATACAGCCAGCA 798
 QY 181 GCGGTGATTGAGACACTGTGCTGCTGAGGCTGAGTCCGCTGCTGTATAC 240
 DB 799 GCGGTGATTGAGACACTGTGCTGCTGAGGCTGAGTCCGCTGCTGTATAC 858
 QY 241 ATCTACTCACTCAGATGAAGTAGCTGACACTGCTGAGGCTGAGTTCAGTTC 300
 DB 859 ATCTACTCACTCAGATGAAGTAGCTGACACTGCTGAGGCTGAGTTCAGTTC 918
 QY 301 GCTTGGAGGGCGAGTCAAGAGATGATCTTGTGTATATGACCGTGTGTAACTG 360
 DB 919 GCTTGGAGGGCGAGTCAAGAGATGATCTTGTGTATATGACCGTGTGTAACTG 978
 QY 361 GATGAGTGGAGGCGACATGATCTGATGATGAGGAGAGATTAACCATGGTTAT 420
 DB 979 GATGAGTGGAGGCGACATGATCTGATGATGAGGAGAGATTAACCATGGTTAT 1038
 QY 421 AAGAAGTATCCAAACGCTTTTAAAGATCCGAGCATTTGGAAGAGAGGCTGACTGT 480
 DB 1039 AAGAAGTATCCAAACGCTTTTAAAGATCCGAGCATTTGGAAGAGAGGCTGACTGT 1098
 QY 481 GTTACAGAGCTGTATCACTCTCAAGCTGGAAAGCTCTGTTCGCGCATGACGTC 540
 DB 1099 GTTACAGAGCTGTATCACTCTCAAGCTGGAAAGCTCTGTTCGCGCATGACGTC 1158
 QY 541 AATGATCTGTACCAACAGAGTTGATTAATCTGCTGCGGAGATTCATTG 600

DB 1159 AATGATCTGTACCAACAGAGTTGATTAATCTGCTGCGGAGATTCATTG 1218
 QY 601 GATGAGTGGAGAGGACACAGATGTGATTTGTTGGGAAACAAGTGTGTGTGC 660
 DB 1219 GATGAGTGGAGAGGACACAGATGTGATTTGTTGGGAAACAAGTGTGTGTGC 1278
 QY 661 TATGTTGAGTATGAGCAAGGCTGCTGTCTCTCAAAAGCTTTGGAGCAATTTCTAC 720
 DB 1279 TATGTTGAGTATGAGCAAGGCTGCTGTCTCTCAAAAGCTTTGGAGCAATTTCTAC 1338
 QY 721 ATTACCGAATGCAACCCATCTGTCTGACAGGCTCATGATGAGTGGTGGTA 780
 DB 1339 ATTACCGAATGCAACCCATCTGTCTGACAGGCTCATGATGAGTGGTGGTA 1398
 QY 781 AAGCTAAATGAAGTATCCGCGCAAGTGTGATTAATCTGACAGAGAAATAGAT 840
 DB 1399 AAGCTAAATGAAGTATCCGCGCAAGTGTGATTAATCTGACAGAGAAATAGAT 1458
 QY 841 GTAGTACAGGAGGAGCAATTTGATGATGAAGAAACATTTGATGATGATGAGG 900
 DB 1459 GTAGTACAGGAGGAGCAATTTGATGATGAAGAAACATTTGATGATGATGAGG 1518
 QY 901 CACTCAACACAGAAATGATGTGACAGACCTCCGACCTCGAGCTGACGAGGAGGA 960
 DB 1519 CACTCAACACAGAAATGATGTGACAGACCTCCGACCTCGAGCTGACGAGGAGGA 1578
 QY 961 GTAGTCTCAGGTGAGCAATGTATCTGCGCAGATGAGCAACGAGTTGCTCTGCA 1020
 DB 1579 GTAGTCTCAGGTGAGCAATGTATCTGCGCAGATGAGCAACGAGTTGCTCTGCA 1638
 QY 1021 GAGGTGCTGACCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 DB 1639 GAGGTGCTGACCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1698
 QY 1081 GCCACACACAGGCTTTGGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 DB 1699 GCCACACACAGGCTTTGGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1758
 QY 1141 CAGGATGTGATCTGCTCTTAAAGAAATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 1759 CAGGATGTGATCTGCTCTTAAAGAAATGATGATGATGATGATGATGATGATGATGAT 1818
 QY 1201 TCATTTGATGCGCACTTACAGAGGTGACAGATGACAGCAAGAAATATCTGGAGTCAAC 1260
 DB 1819 TCATTTGATGCGCACTTACAGAGGTGACAGATGACAGCAAGAAATATCTGGAGTCAAC 1878
 QY 1261 AAAAATGGCCATCAAACTAATTAATTAAGATAC 1296
 DB 1879 AAAAATGGCCATCAAACTAATTAATTAAGATAC 1914
 RESULT 4
 AAV73924
 ID AAV73924 standard; DNA; 3616 BP.
 XX AAV73924;
 XX 04-MAR-1999. (first entry)
 XX Human SAHH DNA #1.
 XX S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;
 XX treatment; infection; cancer; autoimmune disease; detection; diagnosis;
 XX gene mapping; antisense; therapy; antagonist; immunosay; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX CDS 56..1558
 XX FT /*tag- a
 XX FT /product- "SAHH"
 XX PN US5854023-A.

XX 29-DEC-1998.
 PD 17-JUL-1997; 97US-0896005.
 PF 17-JUL-1997; 97US-0896005.
 PR 17-JUL-1997; 97US-0896005.
 XX (INCY-) INCYTE PHARM INC.
 PA Corley NC, Hillman JL, Lai P, Shah P;
 PI WPI: 1999-094906/08.
 DR P-PSDB; AAM90061.
 XX

PT Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase -
 PT for production of recombinant enzyme, useful for diagnosis,
 PT treatment and prevention of cancers, infections and autoimmune
 PT diseases

PS Disclosure; Fig 1A-I; 40pp; English.

XX This sequence encodes a human S-adenosyl-5-homocysteine hydrolase (SAHH).
 CC The SAHH protein can be used to generate specific antibodies and in drug
 CC screening to identify specific binding agents. Antagonists of the
 CC protein are used to treat or prevent a wide range of viral, bacterial,
 CC (leukemia, lymphoma or solid tumors), and many autoimmune diseases
 CC (e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes
 CC mellitus, multiple sclerosis etc). All these conditions may be treated by
 CC expressing antisense sequences, triplex-forming agents or ribozymes
 CC directed against the nucleic acid. The nucleic acid and its fragments can
 CC be used as probes or primers for detecting and quantifying gene
 CC expression, for diagnosis or monitoring of disease, to identify genetic
 CC variations, mutations or polymorphisms, in gene mapping and as antisense
 CC therapeutics. Antibodies are used directly as antagonists, indirectly to
 CC deliver active agents to SAHH-expressing cells, to diagnose and monitor
 CC diseases in standard immunoassays, in competitive drug screens and to
 CC isolate the protein from natural sources.

XX Sequence 3616 BP; 1017 A; 782 C; 826 G; 991 T; 0 other;

Query Match 99.9%; Score 1294.4; DB 20; Length 3616;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1295; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACAGGCGACAGCAATTTCTGTGTAAGAACATCAAGCAGGCAATTTGGACGCGG 60
 DB 260 TCACAGGCGACAGCAATTTCTGTGTAAGAACATCAAGCAGGCAATTTGGACGCGG 319
 QY 61 GAGATTGAGTTCAGACAGACATGCTGCTGATTTCACTAGAGAAAGCTGCTCAG 120
 DB 320 GAGATTGAGTTCAGACAGACATGCTGCTGATTTCACTAGAGAAAGCTGCTCAG 379
 QY 121 GGGAGAGAGCCCTGGCTGGGCTAAATAGTGGCTGACACATGACAGCCAGACA 180
 DB 380 GGGAGAGAGCCCTGGCTGGGCTAAATAGTGGCTGACACATGACAGCCAGACA 439
 QY 181 GCGGTGTTGATTGAGACACTGTGCTGGGCTGCTGAGCGGCTGCTGCTGTAAC 240
 DB 440 GCGGTGTTGATTGAGACACTGTGCTGGGCTGCTGAGCGGCTGCTGCTGTAAC 499
 QY 241 ATGACTGACTGAGATGAGTACTGCTGACAGCTGCTGAGGCTGAGGTTGCAAGTTC 300
 DB 500 ATGACTGACTGAGATGAGTACTGCTGACAGCTGCTGAGGCTGAGGTTGCAAGTTC 559
 QY 301 GCTTGAAGAGGCGAGTCAAGAGTACTTCTGCTGATTTGACCCGCTGTGTAACATG 360
 DB 560 GCTTGAAGAGGCGAGTCAAGAGTACTTCTGCTGATTTGACCCGCTGTGTAACATG 619
 QY 361 GATGGGTGGGCGGCAATGATGCTGATGATGGGGGAGATTAAACCACTGGGTTTAT 420
 DB 620 GATGGGTGGGCGGCAATGATGCTGATGATGGGGGAGATTAAACCACTGGGTTTAT 679

QY 421 AAGAATATCCAAAGCTGTTTAAAGATCCGAGCATTTGTGGAAGAGCGTCACTGCT 480
 DB 680 AAGAATATCCAAAGCTGTTTAAAGATCCGAGCATTTGTGGAAGAGCGTCACTGCT 739
 QY 481 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGCTGTGTTCCGGCCATGAAAGCTC 540
 DB 740 GTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGCTGTGTTCCGGCCATGAAAGCTC 799
 QY 541 AATGATTTCTGTATCCAAACAGAGTTTAAATACCTGTACTGCTGCCGAGATCCATTG 600
 DB 800 AATGATTTCTGTATCCAAACAGAGTTTAAATACCTGTACTGCTGCCGAGATCCATTG 859
 QY 601 GATGGCTGGAAGAGGACACAGATGATGTTGTTGGGAAACAAAGTGGTGTGTTGGC 660
 DB 860 GATGGCTGGAAGAGGACACAGATGATGTTGTTGGGAAACAAAGTGGTGTGTTGGC 919
 QY 661 TATGGTGAAGTGAAGGAGGAGGCTGTGCTGTCTCAAAAGCTCTTGAGCAATTTGCTAC 720
 DB 920 TATGGTGAAGTGAAGGAGGAGGCTGTGCTGTCTCAAAAGCTCTTGAGCAATTTGCTAC 979
 QY 721 ATTACCGAAATGACACCCATCTGTGCTGTGACAGGCTGATGATGATGATGATGATG 780
 DB 980 ATTACCGAAATGACACCCATCTGTGCTGTGACAGGCTGATGATGATGATGATGATG 1039
 QY 781 AAGCTAAATGAAGTATCGGCAAGTGTGATGCTGATGATGATGATGATGATGATGAT 840
 DB 1040 AAGCTAAATGAAGTATCGGCAAGTGTGATGCTGATGATGATGATGATGATGATGAT 1099
 QY 841 GTATGACAGCGGAGACACATTTGATGATGATGATGATGATGATGATGATGATGATG 900
 DB 1100 GTATGACAGCGGAGACACATTTGATGATGATGATGATGATGATGATGATGATGATG 1159
 QY 901 CACTCCAAACAGAAATGATGTGACACAGCTCCGACCTCCGAGACTGACGTGGAGCGA 960
 DB 1160 CACTCCAAACAGAAATGATGTGACACAGCTCCGACCTCCGAGACTGACGTGGAGCGA 1219
 QY 961 GTACGTTTCAGGTGAGGACCATGATGATGATGATGATGATGATGATGATGATGATG 1020
 DB 1220 GTACGTTTCAGGTGAGGACCATGATGATGATGATGATGATGATGATGATGATGATG 1279
 QY 1021 GAGGCTGCTGATCAATTTGAGTGTGCTCCACAGTTCACCTTTGTTCTGCTCATCA 1080
 DB 1280 GAGGCTGCTGATCAATTTGAGTGTGCTCCACAGTTCACCTTTGTTCTGCTCATCA 1339
 QY 1081 GCCACAAACAGAGCTTTGGCAGTATGAACTTAAATGACACCCGAGGGGCGATACAG 1140
 DB 1340 GCCACAAACAGAGCTTTGGCAGTATGAACTTAAATGACACCCGAGGGGCGATACAG 1399
 QY 1141 CAGGATGTGTACTTGTCTTCTTAAGAAATGATGATGATGATGATGATGATGATGAT 1200
 DB 1400 CAGGATGTGTACTTGTCTTCTTAAGAAATGATGATGATGATGATGATGATGATGAT 1459
 QY 1201 TCATTGATGCGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCA 1260
 DB 1460 TCATTGATGCGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCA 1519
 QY 1261 AAAAAATGGGCCATTCAAACTTAATTTACAGATAC 1296
 DB 1520 AAAAAATGGGCCATTCAAACTTAATTTACAGATAC 1555

RESULT 5
 AAS76216/C
 ID AAS76216 standard; cDNA; 5030 BP.

XX AAS76216;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #12020.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.
 OS
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001MO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 XX
 XX WPI; 2001-639362/73.
 DR P-PSDB; ABG12029.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 PS
 PS Claim 1; SEQ ID NO 12020; 103bp; English.
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SO Sequence 5030 BP; 1432 A; 1137 C; 1178 G; 1283 T; 0 other;
 Query Match 62.98; Score 815.8; DB 23; Length 5030;
 Best Local Similarity 78.88; Pred. No. 1,4e-254;
 Matches 1023; Conservative 0; Mismatches 272; Indels 4; Gaps 4;

QY 301 GCTTGAAGGCGGAGTCAGAGATGACTCTGGTGTGTATGACCGCTGTGACATG 360
 DB 4169 GCGTGAAGGAGAGTGTAGAGATGACTTTTGGTGTGTATGATATGTGTAAATG 4110
 QY 361 GATGGGTGAGGCGCAATGATCTCTGATGATGGGAGACTTAACCACTGGGTTAT 420
 DB 4109 GAGGCGTGGAGCGCAACATGATCTGTGATGATGGAGGGATCTTACCACTGGATTAT 4050
 QY 421 AAGAAATATCCAAAGTGTTTAGAGATCCGAGGCAATTGTGGAAGAGAGCTACTGCT 480
 DB 4049 AAAAGATATCCCAATGTTTAAATAAATCAAGGCGATAGTAGAGAGTGTATCGA 3990
 QY 481 GTTCACAGGCTGTATCAAGCTCTCCAAAGCTGGGAAGCTGTGTCCGGCCATGAGTC 540
 DB 3989 GTTCACAGGCTGTATCAAGCTCTCCAAAGCTGGGAAGCTGTGTCCGGCCATGAGTC 3930
 QY 541 AATGATCTGTATCAACAAAGAAATTTGATTAATCTGTACTCTCCGAGATCCATTG 600
 DB 3929 AATGATCTGTATCAACAAAGAAATTTGATTAATCTGTACTCTCCGAGATCCATTG 3870
 QY 601 GATGCGTGAAGAGACCCACAGATGTGATTTGTGGGAAGCAAGTGTGTGTGCG 660
 DB 3869 GATGCGTGAAGAGACCCACAGATGTGATTTGTGGGAAGCAAGTGTGTGTGCG 3810
 QY 661 TATGTGAGGTAGGCGAGGCGTGTGCTGTCTGTCTCCAAAGCTTGGAGCAATTGTCTAC 720
 DB 3809 TATGTGAGGTAGGCGAGGCGTGTGCTGTCTGTCTCCAAAGCTTGGAGCAATTGTCTAC 3730
 QY 721 ATTACCGAAATCGACCCCATCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 780
 DB 3749 GTACTGTAAATTTGACCCCATCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 3690
 QY 781 AAGTAAATGATGATCTCCGCAAG-TCGATGTGTATTAATCTGTACAGAGAAATAGAA 839
 DB 3689 AATTAATGATGATCTCCGCAAGTGTGATGTATTAATCTGTACAGAGAAATAGAA 3630
 QY 840 TGTGTGACAGGAGGAGCACTGTGATGATGATGATGATGATGATGATGATGATGATG 899
 DB 3629 TGTGTGACAGGAGGAGCACTGTGATGATGATGATGATGATGATGATGATGATGATG 3570
 QY 900 CCATCCCAACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 958
 DB 3569 ACATTCACACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3510
 QY 959 GAGTACGTTCTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1018
 DB 3509 GAGTACGTTCTCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3450
 QY 1019 CAGAGGCTGTCTCTCAATTTGAGCTGTCTCTCAAGTTCACCACTTGTGTCTCATCA 1078
 DB 3449 CAGAGGCTGTCTCTCAATTTGAGCTGTCTCTCAAGTTCACCACTTGTGTCTCATCA 3390
 QY 1079 CAGGCAACACAC-AGGCTTGTGGCACTGTATGATGATGATGATGATGATGATGATGATG 1137
 DB 3389 CTGCTACTACTCAAGGCTGTCTGTGATGATGATGATGATGATGATGATGATGATGATG 3330
 QY 1138 AAGCAGATGTGTACTCTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATG 1196
 DB 3329 AAGCAGATGTGTACTCTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATG 3270
 QY 1197 GCCATCATTTATGCGCCACTTACAGAGCTGACAGATGATGATGATGATGATGATGATGATG 1256
 DB 3269 GCTTACCTTTATGCGCCACTTACAGAGCTGACAGATGATGATGATGATGATGATGATGATG 3210
 QY 1257 CAACAAATATGGGCGATCAACACTAATTAATAGATA 1295
 DB 3209 CAATAGATATGGGCGCTTCAAGCTTAATTAATAGATA 3171

RESULT 6
 ABL03693
 ID ABL03693 standard; cdna; 1566 BP.

[illegible]

Db	565	TTTGCGCTGGGCGGGCGAGGAGGAAAGGAACTTGTGGTGGCATGATGCGTGGCTTAAC	624
Qy	358	ATGATGGGTGGCAGGCGCAACATGATCTTGGATGATGGGGGAGACTTAACCATCTGGGTT	417
Db	625	GCCGAGAACTGGCAGCCAAACATGATCTCGGAGCATGGGCGGATGCCACGCACTTGATG	684
Qy	418	TATTAAGAAGTATCCAAACGCTTTAAGAAATCCGAGGCAATTGTGGAAAGAGCGTGACT	477
Db	665	CTAAAAAAGTACCCGACCATCTTCAAGTGTGGTAAAGGAATCGTGGAGAGAGCGTACGC	744
Qy	478	.GGTGTACACAGGCTGATTTGAGCTCTCCAAAGCTGGGAACTCTGTGTTCGGGCAATGAC	537
Db	745	GGAGTGCACCGCTCTCTATCAGCTCTCTAAGGCCGGCAAGCTGACATGCCGCGATGAT	804
Qy	538	GTCAAATGATTCGTATTACCAACAGAAAGTTGATTAAGTCTACTGCTCCGAGAAATCGATT	597
Db	805	GTCAAACGATTCGGTGACAAAGAACAAATTGACAACTCTACAGCTGCAGAAATCCATAC	864
Qy	558	TTGAGTGGCCCTGAAGAGAGACACAGATGTATGTTTGGTGGGAAACAAATGGTGGTGT	657
Db	865	CTGGACACCTTCACAGCCCTCCACGGAGCTATTTTGGTGGCAAGAGAGGCTCTGGTCTGT	924
Qy	658	GACATGATGTGAGTAAGGCAAGGGGCTGTGTGCTGTCTCTCAAAAGCTTTGGAGCAATTGNC	717
Db	925	GGCTACGGCGATGTGGCGAAAGGATGTGCTCTCAGGCTCTTAAGGGGAGAGGCTGCATTGG	984
Qy	718	TACATTTACCGAAATCGAACCCATCTGTGCTCTCAGAGCCTGCATGGATGGGTTCAGGGTG	777
Db	985	TACATCACAGAGATCGATCCCATTTGGGCCCTGCAGAGCCAGTATGGATGGCTTCGGTGTG	1044
Qy	778	GTAAGCTAAATGAATGATCTCCGCAAGTCGATGTGTAATACTTGCCACAGAAATAG	837
Db	1045	GTCAGAGCTGAACGAGGTGATCCCGAAGTGAAGATTTGGTGTGGTGCAGCGCACCGCAACAG	1104
Qy	838	AATGTAAGGACAGGGGAGCACTTGGATCCGATGAATAAACAGTTGATCGTATGCAATTTG	897
Db	1105	AACGGGTGGTGGCGGAGCAACATGGACAGATGAAGATGAAGTGGCTGCATCGTATGCAACATG	1166
Qy	898	GGCCACTCCAAACAGAGAATCGATGTGACCGAGCCTCCGACTCCGAGAGCTGACGTGGAG	957
Db	1165	GGCCACTCGAACACGGAATCGATGTGAATGGCCTCGCAACCCGAGACTTGACCTGGGAG	1224
Qy	958	CGATGACGTTTTCAGGTGGAGCAATGTCTATGTGGCCAGATGGCAACAGATTTCTCTCTG	1017
Db	1225	AAGGTGGCCCTCCAGGTGGATCACATTAATTCGCGCGAGGCAAGTACATCAATCTTCTGTG	1284
Qy	1018	GCAGAGGGTGTCTACTCAATTTGAGCTGTCTCACAGTTCCCAACCTTTGTGTGTCATC	1077
Db	1285	GCCGAGGGCAGGCTGTGTCAATCTGAGTGTCTCCAGCATCCCTCGTTGCCGTGTCTATTC	1344
Qy	1078	ACAGCACAACACAGGCTTTGGCACTGATGAACTGATATGACACCCGAGGGGCGAATC	1137
Db	1345	ACGTGGGCAACCCAGGCGCCGTGGCCGTGATGAGCACTTTTCAATGCCCGCCGCGAGGCTAC	1404
Qy	1138	AAGCAGATGTGATCTGCTTCTTAAGAAATGATGAATAGCTGGCCAGCTTGATGTTG	1197
Db	1405	AAGTGGGATGTCTACTTGTGCTGCCAAGAAATGAGAGATGATGTGGCCAGCCTGCACCTG	1464
Qy	1138	CCATCATTTGATGGCCACCTTAACAGAGCTGACAGATGACCAGCAAAATATGTGGGACTC	1257
Db	1465	CCCACCTTCAGTGGCCCAATTGACGGAGCTGAGCGAGAGCGAGCCCAAGTACATGGGACTG	1524
Qy	1258	AACAAATATGGGCATTCAAACCTAATTTATTCAGATAC	1296
Db	1525	AACAAGGCGCGGTCTCTTTCAGGCCCAACTACTACGCTAC	1563
RESULT 7			
ABL29787			
ID ABL29787 standard; DNA; 1776 bp.			
XX ABL29787;			
AC ABL29787;			

DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 5558.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;
KX	pharmaceutical; gene; ss.
XX	
OS	Drosophila melanogaster.
PX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US092231.
XX	
PR	23-MAR-2000; 2000US-191637P.
XX	
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PERKE) PE CORP NY.
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI; 2001-6556860/75.
XX	
DR	P-PDB; ABB59589.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
PT	
XX	
PS	Claim 1; SEQ ID NO 5558; 21pp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB1616176-AB130511), expressed DNA sequences (AB101840-AB1616175) and the encoded proteins (ABB57737-ABB72072).
CC	
CC	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
CC	
XX	
SQ	Sequence 4406 BP; 1112 A; 1039 C; 1064 G; 1191 T; 0 other:
	Query Match 37.5%; Score 486.2; DB 23; Length 4406;
	Best Local Similarity 66.9%; Pred. NO. 4.1e-147;
	Matches 742; Conservative 0; Mismatches 303; Indels 64; Gaps 1
OY	252 TCAGATGAAGTAGTGCAGACACTGCTGAGAGCGTGAGTTGCAGTGTTCGTTGGAAAGG 311
DB	2112 TTAGAACAAGTGGCTGCTGCTTGTGGCCGAATCCGGAAATCCCAATCTTGCTTGCGCGG 2053
OY	312 CGAGTCAGAAAGATGACTTCTGTGGTGTATTGACCGCTGTGTAACATGATGGATGGGACA 371
DB	2052 CGAAGCAGGAAGAGGACTTCTGTGGTGACATAATGCGCTGTTAACGCCGAGAACTGGCA 1993
OY	372 GGCCAAACATGATCCTCGATGATGGGGAGACACTTAACCACCTGGGTTTAAAGAATATCC 431
DB	1992 GCCAAACATGATCCTCGAGCATGGGGGATCCACGCACTTGATGCTAAAAAAGTACC 1933
OY	432 AAACGTGTTTGAAGATCCGAGGCAATGTGGAAGAGACGAGTGGGTTCACAGGCT 491
DB	1932 GAAGTGTTCAGTTGGTCAAGGAAATCCTCGAAGAGACGTCACCGGAGTGACCCGTT 1873
OY	492 GTATAGCTCTCCAAGCTGGGAACCTGTGTTCGCGGCATGAAGCATGATGATTTCTGT 551
DB	1872 CATATAGCTCTTAAAGCGCGGCAAGCTGACATGCGCGGAGTAAGTCAACGATTCGGT 1813
OY	552 TACCAAACGAAGTTGATTAAGTCTGCTACTGCTGCCGAGATCAATCTTTGGATGGCCTGAA 611
DB	1812 GACAAAGACCAATTTAGCAACCTCTACAGCTGCAAGGATCATCTCTGACAGCCTCAA 1753
OY	612 GAGGACCAAGATGTGATTTGGTGGGAACAAGTGGGTGTGGGTATAGTGAAGG 671

Db	1752	GCAGTCCACGGACAGCTGATGTTGGTGGCAAGCAGGTCGAGTGTGGTGGCTACGGGAGATG	16933
Qy	672	AGGCAGGGCGTGGCTCTCTCTCA-----	697
Db	1692	GGGCAAGGAGTGGCTCAGGCTCTAAAGGGGCGAGTAAATGGCCACCACCAACCATTAATA	16333
Qy	698	-----AACCTTTGGAGCAATTGTCATATTACG	727
Db	1632	TGATATACCTTTGTCACTTTTCAATCTCCAAATCGAGGGCTGTGCAATTGTGTACATCACAG	15733
Qy	728	AAATCGAACCCCATCTGTCTCTCGAGGCCATGATGGGTTCCAGGTGGTAAAGCTTAA	787
Db	1572	AGATCGATCCCATTTGGGCCCTCGAAGGCCAGTATGGATGGCTTCGTTGGTGAACAGCTGA	15133
Qy	788	ATGAATCATCCGGCAAGTGGATGTGTAACTTGCAACAGGAAATAAAGTATAGTGA	847
Db	1512	ACGAGGTGATCCCGCAAGGTAGATGTGTGTGAGCGCGACCCGGCAACAGAAACGTGGTGG	14533
Qy	848	CACGGGAGCACTGGATCGGATGAAAAACGTGTATGATATGACATATAGGGCCACTCA	907
Db	1452	TGCGGAGACATATGAGAACATGAAAGTGGCTGTGATCGTATGCAACATGGGCGACTGA	13933
Qy	908	ACACGAAATGATGTGACACAGCCCTCCGCACTCCGAGAGCTGACAGTGGAGCGAGTACGTT	967
Db	1392	ACACGGAAATGATGTGAAATGGCCTGCGCACACGGGACTTGAACCTGGAGAAAGTGGCT	13333
Qy	968	CTCAGGTGACCATGTATCTGTGGCCAGANTGGCAACAGATGTGTCTCTGCGAGAGGCTC	1027
Db	1332	CCGAGGTGATCACATATATCTGGCGGAGGGCAAGTACATATCTTCTTGGCCGAGGGCA	12733
Qy	1028	GTCATCTCAATTGATGCTGTCTCCAGATGCCACCTTTGTGTCTGCATCACAGGCCAA	1087
Db	1272	GCGTGTCAATCTAGCTGTCTCCAGCATCCCTCGTTCGCCGTCTATCACCTGGGCA	12133
Qy	1088	CACAGGCTTTGGCAGTATAGAACTATATATCCACCCAGGGCGCATCAAGCAGGATG	1147
Db	1212	CCGAGGCCCTGGCCCTGATGAGGCTTTCAATGCCCGGCCGAGGACGTCAAGTGGGATG	11533
Qy	1148	TGTACTGCTTCTTAAGAAATGGAGATATGCTGCCAGCTGTGATCTGCATATTTG	1207
Db	1152	TCTACTGCTCTCCCAAGAGATGGACAGATATGTTCCAGCCTGCGACCTGCCACCTTCG	10933
Qy	1208	ATGGCCACCTTACAGAGCTGACAGATGACCAACAAATATCTGGGACATCAACAAATAG	1267
Db	1092	ATGCCCATTTTACGAGAGCTGAGCAGCAGCAGGCCAATGATGAGTGGACTGAAACAAGGCCG	10333
Qy	1268	GGCCATTCAAACTAATTATTACAGATAC	1296
Db	1032	GTTCCTTCAAGCCCACTACTACCGCTAC	1004
RESULT 9			
ABLT/9045			
ID	ABLT/9045	standard; cDNA; 553 bp.	
XX	ABLT/9045;		
XX	17-MAY-2002 (first entry)		
XX	Human ovarian cancer related cDNA clone SEQ ID NO:2023.		
XX	Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.		
XX	Homo sapiens.		
XX	MO200192581-A2.		
XX	06-DEC-2001.		
XX	29-MAY-2001; 2001MO-US17756.		
XX	26-MAY-2000; 2000US-207484P.		
XX			

PA (CORI-) CORIXA CORP.
 XX
 PI Algate PA, Harlocker SL, Jones R;
 DR WPI: 2002-122075/16.
 XX
 PT Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PT polypeptide
 PS Claim 1; SEQ ID 2023; 489pp; English.
 XX
 CC The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumor
 CC polypeptide encoded by a polynucleotide (III) having a CDNA sequence
 CC (S1) from the 10912 nucleotide sequences as given in AB177023 to
 CC AB187934, (III) encoding (II) having a sequence (S2), a T cell
 CC population of (II), or antigen presenting cells that express (II).
 CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridizes to
 CC (S1) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridizing to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridizing to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC (III) and/or (II) is useful for stimulating and/or expanding T cells
 CC specific for an ovarian tumor protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumor polypeptides
 CC and proteins in tumor cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumor CDNA library using well known
 CC techniques.
 XX
 SQ Sequence 553 BP; 159 A; 136 C; 128 G; 127 T; 3 other:
 Query Match 35.6%; Score 461.4; DB 24; Length 553;
 Best Local Similarity 97.3%; Pred. No. 1.4e-139;
 Matches 510; Conservative 0; Mismatches 9; Indels 5; Gaps 4;
 OY 774 GGTGTAAGCTAATGATGATCGGCAAGTGTGCTAATTAATTCACAGGAAA 833
 DB 1 GGTGTAAGCTAATGATGATCGGCAAGTGTGCTAATTAATTCACAGGAAA 60
 OY 834 TAAGATGTAGTACAGCGGAGCACTTGATCGATGCAAAAAAGTTGATCGATGCA 893
 DB 61 TAAGATGTAGTACAGCGGAGCACTTGATCGATGCAAAAAAGTTGATCGATGCA 120
 OY 894 TATGGCCACTCCACACAGAGAAATGATGATGATGATGATGATGATGATGATG 953
 DB 121 TATGGCCACTCCACACAGAGAAATGATGATGATGATGATGATGATGATGATG 180
 OY 954 GGAGCGAGTACGCTCTGAGTGGAGCATGATGATGATGATGATGATGATGATGATG 1013
 DB 181 GGAGCGAGTACGCTCTGAGTGGAGCATGATGATGATGATGATGATGATGATGATG 239
 OY 1014 CCTGGCAGAGGGTGTCTACTCAATTTGAGTGTCTCCACAGTCCACCTTGTCTGTC 1073
 DB 240 CCTGGCAGAGGGTGTCTACTCAATTTGAGTGTCTCCACAGTCCACCTTGTCTGTC 299
 OY 1074 CATGACAGCCACACACAGGTTTGGCAGTATGATGATGATGATGATGATGATGATG 1133
 DB 300 CATGACAGCCACACACAGGTTTGGCAGTATGATGATGATGATGATGATGATGATG 359
 OY 1134 ATACAGAGAGATGTGATCTGCTTCTCTAAGAAA-TGATGATGATGATGATGATGATG 1192
 DB 360 ATACAGAGAGATGTGATCTGCTTCTCTAAGAAA-TGATGATGATGATGATGATGATG 419
 OY 1193 ATGTCGATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1252
 DB 420 ATGTCGATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 478

OY 1253 GACTCACAATAATGGCCATTAACCTTAATTAATTAATTAATTAATTAATTAATTA 1296
 DB 479 GACTCACAATAATGGCCATTAACCTTAATTAATTAATTAATTAATTAATTAATTA 520
 RESULT 10
 ABV44141/c
 ID ABV44141 standard; CDNA: 504 BP.
 XX
 AC ABV44141;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker CDNA 44132.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN W0200160860-A2.
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 DR WPI: 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 8770; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV6213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 504 BP; 117 A; 150 C; 116 G; 121 T; 0 other:
 Query Match 32.6%; Score 422; DB 23; Length 504;
 Best Local Similarity 100.0%; Pred. No. 9.2e-127;
 Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 158 GTACACATATACAGCCACAGAGGGTGTGATGATGATGATGATGATGATGATGATG 217
 DB 502 GTACACATATACAGCCACAGAGGGTGTGATGATGATGATGATGATGATGATGATG 443
 OY 218 AGTGGCGTGTGCTGTGATCACTCTCACTCAAGTAAGTACTGACAGACTGG 277


```

|||||
DB 3897 GCATATGGACATTCCTGCTCGGAGATGATGTATGCTATTCGGGGTGCATTC 3956
    924 -----GACCAGCCTCCGACCTCCGGAGC 946
    3957 TGGCCTCATATATCAAAAGTTTAAACCATTTAGAAATGGCTTCATACCCCGAGC 4016
    947 TGAGGTGGAGGAGTACGTTCTCAGTGGACCATGTCATCTGCGCAGATGGCAAGAG 1006
    4017 TAACGTGGAGCGTGTCCGTTCTCAAGTACATCAGATGGCGGAGCGAGATGA 4076
    1007 TTGTCCTCCGCGAGAGGTCGTCTACTCAATTTAGCTGCTCCACAGTCCCACTTTG 1066
    4077 TCAATTTGCTCGCGGAGGAGAGTGTGATTTGCTCTCCACCATTTGCTCTTG 4136
    1067 TTCTGTCATACACAGCACAACAGCCTTTGGACATGATGATGATATATACACCG 1126
    4137 TCGATCCGTGGCTCATCCACCCAGCTTTGGCCCTGATGAACTCTTCACAGCC - 4194
    1127 AGGGCGATACAGCAGAGTGTACTGCTCTCTTAAGAAAAGATGATGATGCTTG 1183
    4195 -AGGAGATATAGTCCGATGTACTGCTGCGCAAGAAAATGGTAGGTTCTAG 4250

```

RESULT 12

AA280766/c
ID AA280766 standard; cDNA; 636 BP.

AA280766;

07-APR-2000 (first entry)

Human colon cancer cell line SW480 cDNA clone SEQ ID NO:850.

Human: gene expression product; diagnosis: tumour; colon cancer;
colorectal adenocarcinoma; cell line SW480; cell proliferation;
cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
hyperplasia; ds.

Homo. sapiens.

NO9964576-A2.

16-DEC-1999.

09-JUN-1999; 99MO-IB01062.

10-JUN-1998; 98US-0088801.

(FARB) BAYER CORP.

Endege WO, Steinmann KE, Aetle JH, Burgess CC, Bushnell SE;

PI Carroll E, Catino TJ, Dertl A, Ford DM, Lewis ME, Monahan JE;

Schlegel R;

WPI; 2000-087220/07.

Novel nucleic acids, used to develop products for the diagnosis and

treatment of disorders involving unwanted cell proliferation,

particularly cancers, especially colon cancer

Claim 15; Page 469; 469pp; English.

AA279917 to AA280766 represent double stranded cDNA clones isolated from
the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
cDNA clones can be used to generate antisense oligonucleotides which
can be used for antisense therapy. Methods and products from the present
invention can be used for identifying and/or classifying cancerous cells
present in a human tumour, particularly in solid tumours, e.g.
carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
can be used for developing agents for the diagnosis and treatment of
disorders involving unwanted cell proliferation, such as neoplasia,
dysplasia or hyperplasia.

XX Sequence 636 BP; 156 A; 166 C; 137 G; 147 T; 30 other;
SQ
Query Match 31.4%; Score 406.4; DB 21; Length 636;
Best Local Similarity 99.5%; Pred. No. 1,3e-121;
Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

    159 TACACATCATCACAGCCACAGACAGC-GGTGTTGATTGAGACACTGTGCGGGGCTC 217
    420 TACACATCATCACAGCCACAGACAGCGGTGTTGATTGAGACACTGTGCGGGGCTC 361
    218 AGTCCCGCTGCTGCTGTGTAACATCTACTCACTCACTCAAGAGTGTGCTGACACTGG 277
    360 AGTCCCGCTGCTGCTGTGTAACATCTACTCACTCACTCAAGAGTGTGCTGACACTGG 301
    278 CTGAGGCTGGGTTGTCAGATGCTGCTTGAAGGCGGAGTCAGAGATGACTTCTGTGGT 337
    300 CTGAGGCTGGGTTGTCAGATGCTGCTTGAAGGCGGAGTCAGAGATGACTTCTGTGGT 241
    338 GTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCACATGATCTGATGATGGG 397
    240 GTATTGACCGCTGTGTGAACATGATGGGTGGCAGGCCACATGATCTGATGATGGG 181
    398 GAGACTTAACCCACTGGGTTTATGAAGTATCCAAACGCTTTAAGAGATCCGAGGCA 457
    180 GAGACTTAACCCACTGGGTTTATGAAGTATCCAAACGCTTTAAGAGATCCGAGGCA 121
    458 TTGTGAAGAGAGCGTGTGCTGCTTGCACAGCGTATACACTCCCAACCTGGGAGAGC 517
    120 TTGTGAAGAGAGCGTGTGCTGCTTGCACAGCGTATACACTCCCAACCTGGGAGAGC 61
    518 TGTGTGTTCCGCGCATGACGTCATGATCTGTGTACCAACAGAGTTGATTAACCTGT 577
    60 TGTGTGTTCCGCGCATGACGTCATGATCTGTGTACCAACAGAGTTGATTAACCTGT 1

```

RESULT 13

AACT7823
ID AACT7823 standard; cDNA; 2200 BP.

AACT7823;

08-FEB-2001 (first entry)

Human cancer associated gene sequence SEQ ID NO:217.

Human: cancer associated gene; cancer antigen; detection; cancer;
diagnosis; cytostatic; proliferative; vulnerary; immunomodulatory;
antidiabetic; antiscismatic; antirheumatic; antirachitic; antiviral;
antiinflammatory; antithyroid; antiallergic; antibacterial; cardiatic;
dermatologic; neuroprotective; thrombolytic; coagulant; nootropic;
vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
immune disorder; haematopoietic cell disorder; autoimmune disorder;
allergic reaction; graft versus host disease; organ rejection;
haemostatic; thrombolytic; cardiovascular disorder; infection;
neurological disease; drug screening; ss.

Homo sapiens.

WO200055350-A1.

21-SEP-2000.

08-MAR-2000; 2000MO-US05882.

12-MAR-1999; 99US-0124270.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI; 2000-587533/55.

P-PSDB; AAB43614.

XX Novel, isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer -
XX
PS Claim 1, Page 787-788, 2352pp; English.
XX

CC AAC767607 to AAC784448 encode the human cancer associated proteins given
CC in ABA83398 to ABA84239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vlnetary; immunomodulator;
CC antidiabetic; antiaesthetic; antithrombotic; antitumor;
CC antinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC nototropic; vasotropic; antipsoriatic and antiautogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haemtopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies, and
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and ABA84240 represent sequences used in the exemplification of
CC the present invention.

SQ Sequence 2200 BP; 546 A; 578 C; 607 G; 466 T; 3 other;

Query Match	31.0%	Score 402.4	DB 21	Length 2200
Best Local Similarity	58.4%	Pred. NO. 5.7e-120		
Matches 741	Conservative 0	Mismatches 521	Indels 6	Gaps 2

OY	32	ACATCAGCAGGAGCAAAATTTGGAGCCCGGGAGATTGGATTGCAGAGCAACATGTCG	91
Db	108	ACATCGCCCTGGCTCCCTGGGGAGCCAGGCCCTCGACATTTGCTGAGAACGATGCGG	167
OY	92	CTCTGATTTCACTCAGGAAACGTCGCTCAGGGGGAGAACCCCTTGCTGTGCTAAATAG	151
Db	168	GCTGATCGGATGCGGGAGCGGATACCTGGCGCTCCAAAGCCACTAAGGGCCCGCATCG	227
OY	152	TGGCGTGCACACATCACAAGCCAGACAGCGGTGTGATTGAGACACTCTGTCCCTGG	211
Db	228	CTGGCTGCTCCACATATGACCCCTGGAGAGCGCGCTCTCATTTGAACCCCTGTCACCTGG	287
OY	212	GGGCGACATGGCCGGGCTGCTCTGTAAACATCTACATCACTCAACAATGAAGTACTCAG	271
Db	288	GTGCTGAGGTGACAGTGGCTGACGCTGCAACATCTTCTCACCCAGACCAATGCGCGCTG	347
OY	272	CACCTGGCTGAGAGGCTGGAGATTGCATGTTCCGTTGGAAAGGGGAGTCAAGATATCTCT	331
Db	348	CCATTCGCCAAGGCTGCGCATTTCCGGTGTATGCTCTGGAAAGCGGAACGAGAGGATACC	407
OY	332	GGTGGTGTATTTGACCCGCTGTGTGAACATGATGGGTGGAGGCCAACATGATCTCTGATG	391
Db	408	TGTGGTCATTTGACACAGACCTGTACTTCAAGGAGAGCGGGCCCTCAACATGATTTGGAAG	467
OY	392	ATGGGGGAGACTTAAACCACATGGGTTTATAGAAGTATCCAAACGTGTTTAAAGATCC	451
Db	468	ACGGGGGGACCTTACCAACCTCATCCACACCAAGTACCCGACACTTCTGCAGAGCATCC	527
OY	452	GAGGACTGTGGAGAGAGAGCGTACTGAGTGTTCACAGGCTTATCAGCTGCCAAACGTG	511
Db	528	GAGGCACTCTCGAGAGAACACGCACTGGGGTCCCAACCTTACAAAGTATGATGGCCATG	587
OY	512	GGAAAGCTCTGTGTTCCGGCCATGGAACGTCAATGATTTGTTACCAAAAGAAAGTTGATA	571
Db	588	GGATCTCTCAAGGTGCTGCCCATCATGTGCATGACTCGCTCACCAAGAGCAAGTTTGACA	647
OY	572	ACTGTGTACTGCTGCCCGGAATCCATTTTGGATGGCCTGGAAGAGACCAACAGATGTATGT	631
Db	648	ACCTCTAATGGCTGCGGGAGTCCCTCATAGATGGCAATCAACCGGGCCCAACATATTTGATGA	707

QY	632	TTGGTGGGAAAACAAGTGTGGTGTGGCTATGGTGGAGGTAGGCGAAGGGCGTGTGGCTG	691
.Db	708	TTGCCGGGAAAGTACCGGTGGTAGCAGGCTTATGGTATGTGGGAAAGGGCTGTGCCAGG	767
QY	692	CTCTCAAAAGCTCTTGGAGCAATTGCTACATTACCGAATCGAACCCCATCTGTGCTTGC	751
.Db	768	CCCTCGGGGGTGTGGAGACCCCGCGTCATCATCACCGAGATTGACCCCATCAAGCATTCG	827
QY	752	AGGCTGCATGGATGGGTTCAGGGTGGTAAAGCTTAATGAAGTATATCCGGCAAGTCATG	811
.Db	828	AGGCGCCATGTAGGGGCTATGAGGAGACCACTAGATGAGAGCCCTGTGACAGAGGGCAACA	887
QY	812	TGCTATTAATCTGCACAGGAATAAAGATGTATGAGACCGGGAGCACTTGGATGCGATGA	871
.Db	888	TCTTTGTATCACCACCCACAGAGCTGTATTTAGCATATCTTTGGCCGGCACTTTGACGAGATTGA	947
QY	872	AAACACGTTGATTCGTATGCAATATGGGCACTCCCAACACAGAAATGATGTGACCGACC	931
.Db	948	AGGATGATGCCATTTGTGTGTAACATTTGGACACTTTGACGTGGAATGATGTCAAGTGGC	1007
QY	932	TCCGCACTCCGGAGCTGACGTGGGAGCGAGTACGTTTCAAGTGGACCATGTCAATCTGGC	991
.Db	1008	TCAAGGAGAAAGCCCTGGAGAAAGTGAACATCAAGCCCGAGGTGAGCCGGTATCGGTTGA	1067
QY	992	CAGATGGCAAAAGATTTCTCTCCGCGAGAGGGTGGCTCATTAATTTAGCTGTGCA	1051
.Db	1068	AGAAATGGGCGCCGATCATCTGCTGGCGGAGGGTGGCTGTATACCTGGGTTGTGCCA	1127
QY	1052	CAG--TTCCCAACCTTTGTCTGTCCATCACAGCCACACAAACAGGCTTTGGCACTGTATG	1108
.Db	1128	TGGGGCACCCCGACGTTCTGTATGATTAATCTCTTCACCAACAGAGTGTATGGCGAGATCG	1187
QY	1109	AACCTATATATGCAACCCGAGGGGCGATTCACAGAGATGTACTTGGCTTCTTAAGAAA	1168
.Db	1188	AGCTGTGGACCCATTCACACAAG--TACCCCGTTGGGGTTCATTTCTGCCCCAAGAAGC	1244
QY	1169	TGGATGATTAAGTGTGCCAGCTTGATCTGACATCATATTGTATGCCACCTTACAGAGCTGA	1228
.Db	1245	TGGATGAGGCACTGCTTAAGCCCACTGGGCAACGCTGAATGTGAATGTGACCAAGCTAA	1304
QY	1229	CAGATGACCAAGCAAAATATCTGGGACTCAACAACAAAATGGGCACTTCAAACTAATTAT	1288
.Db	1305	CTGAAGAACCAAGCCAGTACCTGGGCAAGTCTGTGTATGTGGCCCTTCAAGCCGATCACT	1364
QY	1289	ACAGATAC 1296	
.Db	1365	ACCGCTAC 1372	
RESULT 14			
ABV23126			
ID	ABV23126	standard; cDNA; 2658 BP.	
XX	ABV23126;		
XX	ABV23126;		
XX	16-SEP-2002	(first entry)	
XX	Human prostate expression marker	cdNA 23117.	
XX	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;		
XX	pharmacogenomic marker; gene; ss.		
XX	Homo sapiens.		
XX	WO200160860-A2.		
XX	23-AUG-2001.		
XX	20-FEB-2001; 2001MO-US05171.		
XX	17-FEB-2000; 2000US-183319P.		
XX	16-MAR-2000; 2000US-189862P.		

XX 20-FEB-2001; 2001WO-US05171.
 PF 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Schlegel R, Endege WO, Monahan JE;
 DR WPI, 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX Claim 1; Page 6122; 11750pp; English.
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX Sequence 2658 BP; 624 A; 727 C; 730 G; 552 T; 25 other;
 SO
 Query Match 31.0%; Score 402.4; DB 23; Length 2658;
 Best Local Similarity 58.4%; Pred. No. 6,4e-120;
 Matches 741; Conservative 0; Mismatches 521; Indels 6; Gaps 2;
 QY 32 ACATCAAGACGCGAATTTGGACGCGGAGATGAGATTGAGACGACAGACATGCTG 91
 DB 355 ACATCGGCTGGCTGCTGGGAGCGCAAGGCCCTGACATCTGAGAACGAGATGCCGG 414
 QY 92 CTCTGATTTCACAGAAACGCTCCTCAAGGGGAGAAAGCCCTTGGCTGTCTAAATAG 151
 DB 415 GCGCTGATGCTATGCGGGAGCGGTACTGCGCTCCAAAGCCACTGAAGGGCGCCGCACTCG 474
 QY 152 TGGGCTGTACACATCAGACGCCAGAGAGCGGTGATGAGACACTGTGCCCTGG 211
 DB 475 CTGGCTCCCGACATACCGTGAAGAGCGGCCGCTCATTTGAGAACCTCTGCACCTCGG 534
 QY 212 GGGCTCAGTCCCGCTGCTGCTGTTAATCACTCACTCAAGTCAAGTAACTAGCTGAG 271
 DB 535 GTCTGAGGTGAGTGTGCTGAGTCAACATCTCTCCACCAGACCATCGCGGCTG 594
 QY 272 CAGTGGGTGAGGTGAGTGTGAGTGTGCTGAGGAGGAGTCAAGATCACTTCT 331
 DB 595 CCAATGCAAGGCTGGGATTCGCTGATGCTGGAAGGGGAAACGAGACAGAGATAC 654
 QY 332 GGTGTGATTTGACCGCTGTGTGAACATGATGAGTGGTGGAGCCAGCATGATCTGGATG 391
 DB 655 TGTGTGATTTGAGACAGACCTGTACTTCAAGAGCGGGCCCTCAACATGATTTGTGAGCG 714
 QY 392 ATGGGGAGACTTACCCACTGGGTTTAAAGATATCCAAAGCTGTTTAAAGAGATCC 451
 DB 715 ACGGGGGCGACCTACCAACCTCATCCACACAGATACCGGACCTTCTGCCAGGACATCC 774
 QY 452 GAGGCAATGTGGAAGAGGTGACTGTGTACAGAGCTGTATCAGCTCTCCAAAGCTG 511

DB 775 GAGGCAATGTGGAAGAGGTGACTGTGTGTCCACACCTTACAAAGTATGCGCATG 834
 QY 512 GGAAGCTCTGTGTCCGGCATGAACGTCAATGATTTCTGTACCAAAACAGAGTTGTATA 571
 DB 835 GGATTCCTCAAGGTGCTCCATCATATGATGACTCCGTACCAAGAGCAAGTTGTACA 894
 QY 572 ACTGTACTGTGCGGAGATTCATTTTGGATGGCCTGTAAGAGACCAAGATGTGTGT 631
 DB 895 ACCTTATGCTGCGGGAGTCCCTCATAGATGGCATCAACGGGCCACAGATGTATGA 954
 QY 632 TTGGTGGAAACAAAGT 691
 DB 955 TTGCGGCAAGGTAGCGGTGTAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1014
 QY 692 CTCTCAAGCTCTTGGACCAATTTGTCTATACCAAGATGACCCCATCTGTGTGTGTGT 751
 DB 1015 CCTGTGGGGGTTTGGAGCCCGGCTCATCATCACCGAATTTGACCCATCAAGCATGCG 1074
 QY 752 AGGCTGATGAGTGGTGTGAGGTGTAAAGCTAAATGAATGAAATGTCGGCAAGTGTG 811
 DB 1075 AGGCTGATGAGTGGTGTGAGGTGTAAAGCTAAATGAATGAAATGTCGGCAAGTGTG 1134
 QY 812 TCGTAAATCTTGCACAGAAATGAATGTAGTACACAGGAGACTTGTGATGCAATGA 871
 DB 1135 TCTTTGTACACACACAGAGCTGTATGTACATCATCTTGTGGCGGCACTTTAGCAGATGA 1194
 QY 872 AAACAGTTGTATGCTATGCAATATGAGGCACTCCACACAGAAATGATGTGACACACC 931
 DB 1195 AGGATGATGCAATGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1254
 QY 932 TCCGCACTCCGAGCTGACGTGGAGCGAGTACGTTCTCAGTGTGAGACCATGTCTGTGC 991
 DB 1255 TCAACGAGAAACCGGTGAGAGAGGAAATCAATCAACGCCAGGTGAGACCGTATGCGTGA 1314
 QY 992 CAGATGGAACAGAGTTCTCTCTGTGAGAGGCTGCTCTCAATTTGAGCTGCTCA 1051
 DB 1315 AGAATGGGCGCGGATCACTCTGTGTGCGAGGCTGCGTGTCAACCTGGGTTGTGCA 1374
 QY 1052 CAG--TTCCACCTTGTGTCTGTCCATCAACAGCCACAAACAGGCTTTGGCAGCTAGT 1108
 DB 1375 TGGGCAACCCAGCTGTCTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1434
 QY 1109 AACTCTATATGACCCGAGGGGCGATACAAAGACAGATGTACTTGTCTTAAGAAA 1168
 DB 1435 AGCTGTGAGCCATCCAGACAG--TACCCGTTGGGGTTCATTTCTCTGCCAAGAAAGC 1491
 QY 1169 TGGATGAATAGTGTGCAAGCTTGCATCTGCCATCATTTGATGCCACCTTACAGAGCTGA 1228
 DB 1492 TGGATGAGGAGTGTGAGAGCCCACTGGGCAAGCTGAATGTGAAGTGAACAGCTTAA 1551
 QY 1229 CAGATGACCAAGAAATATCTGGGAGCTCAACAAATATGGGCAATCAAACTAATATT 1288
 DB 1552 CTGAGAAAGCAAGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1611
 QY 1289 ACAGATAC 1296
 DB 1612 ACCGCTAC 1619

Search completed: April 21, 2003, 00:05:52
 Job time : 301.45 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 23:48:04 ; Search time 57.932 Seconds
(Without alignments)
6860.687 Million cell updates/sec

Title: US-09-782-051-1_COPY_549_1844

Perfect score: 1296
Sequence: 1 tccaagggcagcagcaattt.....aacctaattatcagatac 1296

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 segs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, NA: *
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
C 1	406.4	31.4	636	US-09-328-111-850	Sequence 850, App
2	314.4	24.3	2211	US-09-318-448-26	Sequence 26, App
3	314.4	24.3	2211	US-09-347-878-2	Sequence 2, App
C 4	236.2	18.2	4411529	US-09-103-840A-1	Sequence 1, App
5	222.6	17.2	1812	US-08-669-536-1	Sequence 1, App
6	214.6	16.6	1767	US-08-930-894-1	Sequence 1, App
7	117	9.0	2226	US-08-896-005-2	Sequence 2, App
8	117	9.0	2226	US-09-347-878-3	Sequence 3, App
9	72.6	5.6	289	US-08-204-740-8	Sequence 8, App
10	72.6	5.6	289	US-09-081-167A-8	Sequence 8, App
11	72.6	5.6	289	US-09-081-395-8	Sequence 8, App
12	72.6	5.6	289	US-09-416-833-8	Sequence 8, App
13	72.6	5.6	289	US-09-416-833-8	Sequence 8, App
14	68.4	5.3	285	US-08-204-740-6	Sequence 6, App
15	68.4	5.3	285	US-09-081-167A-6	Sequence 6, App
16	68.4	5.3	285	US-09-081-395-6	Sequence 6, App
17	68.4	5.3	285	US-09-416-833-6	Sequence 6, App
18	68.4	5.3	285	US-09-416-833-6	Sequence 6, App
19	68.2	5.3	584	US-08-998-416-249	Sequence 249, App
C 20	41.8	3.2	7218	US-08-232-463-14	Sequence 14, App
21	40	3.1	50341	US-08-247-901C-1	Sequence 1, App
22	40	3.1	50341	US-09-075-904-1	Sequence 1, App
23	40	3.1	52297	US-09-426-436-1	Sequence 1, App
24	40	3.1	52297	US-08-705-557-1	Sequence 1, App
25	36	2.8	20137	US-09-262-773-206	Sequence 206, App
26	36	2.8	20138	US-09-262-773-9	Sequence 9, App
27	36	2.8	23071	US-09-262-773-210	Sequence 210, App

C 28	33.6	2.6	495	1	US-08-133-711-41	Sequence 41, App
C 29	33	2.5	50000	4	US-09-146-053-3	Sequence 3, App
30	32.8	2.5	31571	1	US-08-323-443-1	Sequence 1, App
31	32.8	2.5	53526	3	US-08-658-136-2	Sequence 2, App
32	32.8	2.5	53577	3	US-08-658-136-1	Sequence 1, App
33	32.6	2.5	2101	1	US-08-106-761-1	Sequence 1, App
C 34	32.4	2.5	327	1	US-08-652-859-4	Sequence 4, App
C 35	32.4	2.5	327	1	US-08-919-706-4	Sequence 4, App
C 36	32.4	2.5	327	2	US-09-153-751-4	Sequence 4, App
37	31.4	2.4	4973	4	US-09-381-862-6	Sequence 6, App
38	31.2	2.4	233	2	US-08-687-080-108	Sequence 108, App
C 39	31	2.4	3306	4	US-09-770-170-7	Sequence 7, App
C 40	30.8	2.4	1614	4	US-09-522-217-7	Sequence 7, App
C 41	30.8	2.4	2665	3	US-09-040-005-1	Sequence 1, App
42	30.6	2.4	3186	3	US-08-863-102-3	Sequence 3, App
43	30.6	2.4	4049	1	US-08-162-809-17	Sequence 17, App
44	30.6	2.4	4097	1	US-08-162-809-11	Sequence 11, App
C 45	30.2	2.3	678	4	US-09-461-697-414	Sequence 414, App

ALIGNMENTS

RESULT 1
US-09-328-111-850/c
Sequence 850, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Dertl, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 850
LENGTH: 636
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(636)
OTHER INFORMATION: n - A,T,C or G
US-09-328-111-850

Query Match 31.4%; Score 406.4; DB 4; Length 636;
Best Local Similarity 99.5%; Pred.No. 2.5e-120;
Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 159 TACACACATACAGCCAGCAGC-GGTGTGATTGAGACATCTGCGCCCTGGGGGCTC 217
DB 420 TACACACATACAGCCAGCAGCAGCAGCGGTGTGATTGAGACATCTGCTCTGGGGGCTC 361
QY 218 AGTCCGCTGCTGCTGCTTGAACATCTACTCACTAGATGAAGTAGCTGACAGACTGG 277
DB 360 AGTCCGCTGCTGCTGCTTGAACATCTACTCACTAGATGAAGTAGCTGACAGACTGG 301
QY 278 CTAGAGCTGAGTGTGAGTGTGCTTGGTGAAGGCGCAGTCAAGATGACTTGTGTGT 337

Db 300 CTGAGGCTGAGATGTCAGATGTTTCCTTGGAAGGGGAGATCAGAAAGATGATCTGTGCTGT 241
Qy 338 GTATTGACCGCTGTGTGTAACATGATGGGTGGCAGGCCAACATGATCTGTGATGATGGGG 397
Db 240 GTATTGACCGCTGTGTGTAACATGATGGGTGGCAGGCCAACATGATCTGTGATGATGGGG 181
Qy 398 GAGATTAACCCACTGGGTTTATAGAAGTATCCAAACGTTTAAAGAAATCCGAGGCA 457
Db 180 GAGACTTAACCCACTGGGTTTATAGAAGTATCCAAACGTTTAAAGAAATCCGAGGCA 121
Qy 458 TTGTGAAGAGAGCGTGAAGTGTGTTCACAGCTGTATCAGCTCTCCAAAGCTGGGAAGC 517
Db 120 TTGTGAAGAGAGCGTGAAGTGTGTTCACAGCTGTATCAGCTCTCCAAAGCTGGGAAGC 61
Qy 518 TCTGTGTCGGCCATGAGCTCAATGTTCTGTACCAACAGAAAGTTGATTAATCTGT 577
Db 60 TCTGTGTCGGCCATGAGCTCAATGTTCTGTACCAACAGAAAGTTGATTAATCTGT 1

RESULT 2

US-09-318-448-26
; Sequence 26, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Steinhilber, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-26

Query Match 24.3%; Score 314.4; DB 4; Length 2211;
Best Local Similarity 61.9%; Pred. No. 1,6e-90;

Matches 498; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

Qy 32 ACATCAAGCAGCAGAAATTTGAGCGCGGAGATTGAGATTGACAGAGCAGACATGCTG 91
Db 79 ACATCGGCGCTGCTGCTGCGGAGCGCAAGGCCCTGACATTGCTGAGAGACGATGCCGG 138
Qy 92 CTCGATTTCTACAGAAAGCTGCTCAGGGGAGAGGCCCTTGCTGCTTAATAG 151
Db 139 GCCGTGATGCGTATCGGGAGCGTACTCGGCTCCAAAGCCCTCAAGGGCGCCGCAATCG 198
Qy 152 TGGCTGACACATCAGACAGCCAGACAGCGGTTGATTGAGACACTGTGCCCTGG 211
Db 199 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 258
Qy 212 GGGCTCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 271
Db 259 GTGCTGAGGTCAGTGGTCCAGCTGCAACATCTTCCACCAACCAACATGCGGCGTG 318
Qy 272 CACTGCTGAGGTCAGTGGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 331
Db 319 CCAATGCGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 378
Qy 332 GGTGCTGATGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
Db 379 TGTGCTGATGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 438
Qy 392 ATGGGAGAGATTAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451
Db 439 ACGGGGAGAGCTCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 498
Qy 452 GAGCATTTGTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 511

Db 499 GAGGATCTCTGAGGAGACACAGACTGGGGTCCAAACCTCTACAAAGTATGAGCAATG 558
Qy 512 GGAAGCTCTGTGTCGCGCCATGAACCTGATGATTTGTTACCAAGAGATTGATA 571
Db 559 GGAAGCTCTGTGTCGCGCCATGAACCTGATGATTTGTTACCAAGAGATTGATA 618
Qy 572 ACTGTGACTGCTGCGGAGATTCATTTGATGATGCGCTTAAGAGACCAACATGATGAT 631
Db 619 ACCCTATGAGCTGCGGAGTCCCTCATATGATGATGATGATGATGATGATGATGATGATG 678
Qy 632 TTGCTGGAAGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 691
Db 679 TTGCTGGAAGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 738
Qy 692 CTCGAAAGCTCTTGAAGCAATGCTTACATTAACCAATGACCAATGCTGCTGCTGCTGCTGCTGCTG 751
Db 739 CCGTGGGAGGTTTGGAGACCGCGCATCATCAGCAGATGATGATGATGATGATGATGATGATGATG 798
Qy 752 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 811
Db 799 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 858
Qy 812 TCGTATTAATCTTGCACAGAGATA 835
Db 859 TCTTTGACCAACACAGAGCTGTA 882

RESULT 3

US-09-347-878-2
; Sequence 2, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human S-adenosylhomocysteine hydrolase cDNA
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: M61831/GenBank
US-09-347-878-2

Query Match 24.3%; Score 314.4; DB 4; Length 2211;
Best Local Similarity 61.9%; Pred. No. 1,6e-90;

Matches 498; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

Qy 32 ACATCAAGCAGCAGAAATTTGAGCGCGGAGATTGAGATTGACAGAGCAGACATGCTG 91
Db 79 ACATCGGCGCTGCTGCTGCGGAGCGCAAGGCCCTGACATTGCTGAGAGACGATGCCGG 138
Qy 92 CTCGATTTCTACAGAAAGCTGCTCAGGGGAGAGGCCCTTGCTGCTTAATAG 151
Db 139 GCCGTGATGCGTATCGGGAGCGTACTCGGCTCCAAAGCCCTCAAGGGCGCCGCAATCG 198
Qy 152 TGGCTGACACATCAGACAGCCAGACAGCGGTTGATTGAGACACTGTGCCCTGG 211
Db 199 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 258
Qy 212 GGGCTCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 271
Db 259 GTGCTGAGGTCAGTGGTCCAGCTGCAACATCTTCCACCAACCAACATGCGGCGTG 318
Qy 272 CACTGCTGAGGTCAGTGGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 331
Db 319 CCAATGCGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 378

```

Oy 332 GGTGGTGTATTTAGACCGCTGTGTGAACATGATGGATGGGTGCAGGCCAACATGATCTGGATG 391
Db 379 TGTGTGCATTGAGCAGACCCCTGTACTTCAAGAGAGGGCCCCCTCAACATGATTTCTGACG 438
Oy 392 ATGGGGGAGACTTAACCCACTGGGTTTATTAAGAAGTATCCAAACGCTTTTAAGAAGATCC 451
Db 439 ACGGGGGGAGACTTCCCAACTCATCCACACCAAGTACCCGACAGCTTCTGCCAGGCATCC 498
Oy 452 GAGGATGTGTGAAGAGAGCGGTGACTGGTGTTCACAGGCTGTATCAGCTGTCCAAAGCTG 511
Db 499 GAGGATCTTCTGAGAGACCAAGACTGGGGTCCACAACTCTTCAAGTATGATGGCCAATG 558
Oy 512 GGAACTCTGTGTTCGGCCATGAAACGTCATGATTTCTGTTACCAACAGAAATTTGTATA 571
Db 559 GGATCTCAAGGTGGCTCTCCATCATGTCAATGATCCGTGACCAAGAGCAAGTTTGACA 618
Oy 572 ACTGTACTGTCTCCGGAATATCATTTTGGATGGGCTTAAGAAGACCAACATATGATGT 631
Db 619 AACTGTATGGCTCTCCGGAGTCCCTCATATGATGATGATGATGAGGGGCCACAGATGTGATGA 678
Oy 632 TTGGTGGGAAACAATGTGTGTGTGTGTGGCTATGTGTGAGGTAGGCAAGGGCTGTGTGCTG 691
Db 679 TTGCCGGCAAGGTACGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 738
Oy 692 CTCTCAAAAGCTCTTGGAGCAATTTCTCTACATTTACCGAAATGACCCCATCTGTGCTGCG 751
Db 739 CCTTCGGGGGTTTCGAGACCCCGCTCATCATCAGATGACACGAGATGACCCCATCAAGCATGCG 798
Oy 752 AGGCTGCATGGATGGGTTCACAGGCGGTGAAGCTTAATGAATGATATCCGGCAAGTCGATG 811
Db 799 AGCGTCGCATGTAGAGGCGTATGAGTGTGACCAACCATGTGATGAGGCCCTGTACAGAGGGCAACA 858
Oy 812 TCGTAAATACTTGACACAGGAATA 835
Db 859 TCTTTGTACACACACAGGCTGTA 882

RESULT 4
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 2436-2007_00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 441529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

Query Match	Similarity	18.2%	Score 236.2	DB 4:	Length 4411529
Best Local	Similarity	56.6%	Pred. No. 1.4e-63		
Matches	Conservative	0	Mismatches 375	Indels 9	Gaps 3
Qy	422	AGAACTATCCAAAGCTGTTTAGAAGATCCAGCATTTGTGAGAGAGACCTGACTGGG	481		
Db	3629036	ACAAATGGACCAAGTATAGCCAGATGCTGCAAGGGCGTCACCGAGAGACCAACACGGCG	3628977		
Qy	482	TTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGAGCTGTGTTCGCGGCATGAAGTCA	541		
Db	3628976	TGCGCGGCGCTTACCAATATCCCGCGGCGCGGAGATCGCTTCCCGCGATCAACGCA	3628917		
Qy	542	ATGATTTCTGTACCAACAGAAATTTGATTAATCTGTACTGCTGCCGAGATTCATTTTGG	601		

Db 3628916 ACGAGCTGGTGACCAGATCCAAATTGCAACAAGTACGGCACCTCGGCACTCCCTGTATCG 3628857
OY 602 ATGGCTTAAGAGGACCACCAAGATGTGATGTTTGGTGGGAAACAAAGTGCGTGTGGTGGCT 661
Db 3628856 ACGGCATCAACCCGGGGACCGAGCCGGTGTATGGCGGGTAAGAAAGTCTCATCTGCGGGCT 3628797
OY 662 ATGGTGAGTGAAGGCAAGGGCTGTGCTGTCTCTCAAAAGCTCTTGAGACAAATTGTCTCA 721
Db 3628796 ACGGCGACGTCGTGGTAAGGGCTGTCCGGAGGCGATGAAGAAAGGCGCAGGAGCCCGGGTCTCCG 3628737
OY 722 TTACCGAAATCGACCCCATCTGTGTCTGTGCAGAGCCCTGCATGAGATGGTTCAGGGTGGTAA 781
Db 3628736 TCACCGGAGATCGACCCCATCAACCCGCTGTGACGGCCATGAATGAGGGCTTGTGACCTGTCTCA 3628677
OY 782 AGCTAAAGAAGTACTCCGGGCAAGTCGATGTGCTAATTAATCTGCACAGAAATAGAATG 841
Db 3628676 CCGTTCGAGAGGCCATGGGGGACCCGACATCTGTATACCGGACCGGCACAAABACA 3628617
OY 842 TAGTGACACAGGGAGCACTTGGATCGCATGAAAAACAGTGTATCGTATGCAATATGGGCC 901
Db 3628616 TCATCAATGCTCGACACATTTAAGGCGATGAAGAGCACACGGCATCTGGGAAATATCGGCC 3628557
OY 902 ACTCCAAACAAGAAATTCATGTGACACACCTCCGACATCCGGAAGCTGACATGGGAGCGCAG 961
Db 3628556 ACTTCGACAACGAATGACATGACATGGCCGGGCTCGGGCGCCTCGGGGCAACAGGGTCAACG 3628497
OY 962 TACGTTCCAGGTGAGACCATGTCACTGTGGCCAGA---TGCCAAAGAGATTGTCCTCTGG 1018
Db 3628496 TCAAGCTCAGAGTGCACCTGTGGACCTTTGGCGACAGGGCCGCTGCATCATGCTGCTGT 3628437
OY 1019 CAGAGGGTGTCTACTCAATTTGAGCTGTCCACAGTTCCACG---TTTGTCTGTCCA 1075
Db 3628436 CCGAGGGGCGGCTCTCAACCTGTGGCAATGCCACCGGGCACCCCTCGTCTGTATGAGCA 3628377
OY 1076 TCACAGCCACAACAAGGCTTTGGACAGATGTGAAGCTATATATACACCOCGAGGGGCGAT 1135
Db 3628376 ACAGCTTGCGTAAACAGACGATGCCGCCAGATCGAGCTGTGACCACAGACGACAG---T 3628320
OY 1136 ACAGCAGGATGTACTTGTCTCTTAAGAAATGATGAATACGTTCCAGCTTGCAATC 1195
Db 3628319 ACGACAAGAGAGTGTACCGGCTGTGCCAAGCAGCTCGACGAGAGAGAGTGTGATCATG 3628260
OY 1196 TCCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAACCAAAATATCTGGGAC 1255
Db 3628259 TCGAGGCGCTTGGGGGTCACTCTGACCAAGCTGACCAAGAGACGCGAAATACCTGGGCG 3628200
OY 1256 TCAACAAAATGGGCAATTCAAACCTAATTATTAAGATATC 1296
Db 3628199 TCGAGCTGGAAGGTCCCTTACAGACCGGAGACACTACCGCTAC 3628159

RESULT 5
 US-08-669-536-1
 : Sequence 1, Application US/08669536
 : Patent No. 5910444
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: MASUTA, CHIKARA
 : APPLICANT: MASUTA, KYOKO
 : APPLICANT: BEHARA, YEHARA
 : APPLICANT: TANAKA, HIDEO
 : APPLICANT: KUMATA, SHIGERU
 : TITLE OF INVENTION: ORGANISMS IN WHICH THE EXPRESSION OF
 : TITLE OF INVENTION: S-ADENOSYLHOMOCYSTEINE HYDROLASE GENE IS
 : NUMBER OF SEQUENCES: 4
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
 : STREET: PO BOX 747
 : CITY: FALLS CHURCH
 : STATE: VA
 : COUNTRY: USA
 : ZIP: 22040-0747
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/669,536
 FILING DATE:
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: MURPHY JR, GERALD M
 REGISTRATION NUMBER: 28,977
 REFERENCE/DOCKET NUMBER: 1254-128
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1812 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 US-08-669-536-1

Query Match 17.2%; Score 222.6; DB 2; Length 1812;
 Best Local Similarity 56.4%; Pred. No. 3.7e-61;
 Matches 499; Conservative 0; Mismatches 374; Indels 12; Gaps 4;

DB 424 AAGTATCCAAAGCTGTTTAAAGATCGAGGCAATGTGGAAGAGAGGCTGCTGTT 483
 |||||
 DB 634 AATATACCAAGATGAAGAAAGACTGTCGTTGTTCTGAGAAATACACAGTGAATT 693
 |||||
 DB 484 CACAGGCTGTACAGCTCTCCAAAGCTGGGAAGCTGTGTCGGCCATGACGTAAT 543
 |||||
 DB 694 AAGAGGCTTATTCAGATGACGCTAATGGAATCTTCTTCCGCTATTAACTTTAT 753
 |||||
 DB 544 GATCTGTACCAAGAAAGATTGATGATGATGATGATGATGATGATGATGATGAT 603
 |||||
 DB 754 GATCTGTACCAAGAAAGATTGATGATGATGATGATGATGATGATGATGATGAT 813
 |||||
 DB 604 GCGCTGAAGAGGACACAGATGATGATGATGATGATGATGATGATGATGATGAT 663
 |||||
 DB 814 GGTTCATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 873
 |||||
 DB 664 GGTGAGGAGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
 |||||
 DB 874 GGAATGTGCGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 933
 |||||
 DB 724 ACCGAATCGACCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783
 |||||
 DB 934 ACCGAGATGACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 993
 |||||
 DB 784 CTAATGATGATCGGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 843
 |||||
 DB 994 CTAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1053
 |||||
 DB 844 GTGACAGGAGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 903
 |||||
 DB 1054 ATCATGTTGACACATGAGAGAAATGAAAGAAATGCAATGCTGTTGCAATGTTG 1113
 |||||
 DB 904 TCCAAACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 |||||
 DB 1114 TTTGCAACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173
 |||||
 DB 961 GTACGTTCTAGGTGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
 |||||
 DB 1174 ATTAAAGCTCAAAACGACATGATGATGATGATGATGATGATGATGATGATGAT 1233
 |||||
 DB 1018 GCGAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1074
 |||||
 DB 1234 GCTAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1293
 |||||
 DB 1075 ATCAACCCACAAACAGAGCTTTGGCACTGATGATGATGATGATGATGATGATGAT 1131
 |||||

DB 1294 TGCTGCTCATTACCAAGTCAATTCGCCCACTGCGACTTGTGGAATGAAAGCACTGG 1353
 |||||
 DB 1132 CGATCAAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1191
 |||||
 DB 1354 AAGTATGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1413
 |||||
 DB 1192 CATCTGCATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1251
 |||||
 DB 1414 CATCTGCAGAAAGCTGCGAGCCAAAGCTTACCAAACTTTCGAGAGATCAAGCTGAT 1473
 |||||
 DB 1252 GGACACCAAAAAATGGCCATTCACCTTATTTATACAGTAC 1296
 |||||
 DB 1474 AGCGTTCAGTTGAGGCTCTTACCAAGCTGCTCACTACAGGTAC 1518
 |||||

RESULT 6 US-08-930-894-1

Sequence 1, Application US/08930894
 Patent No. 6037524
 GENERAL INFORMATION:
 APPLICANT: GREENLAND, Andrew James
 APPLICANT: DRAPER, John
 APPLICANT: SKIPSEY, Marc
 APPLICANT: WARNER, Simon
 TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEIN HYDROLASE PROMOTER
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Pillsbury Madison & Sutro
 STREET: 1100 New York Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005-3918
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: MS word
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/930,894
 FILING DATE: 09-OCT-1997
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/GB96/00882
 FILING DATE: 10-APR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9507381.3
 FILING DATE: 10-APR-1995
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1767 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: SHH GENE FROM ASPARAGUS
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 26..1483
 OTHER INFORMATION: /codon_start= 26
 US-08-930-894-1

Query Match 16.6%; Score 214.6; DB 3; Length 1767;
 Best Local Similarity 55.8%; Pred. No. 1.3e-58;
 Matches 494; Conservative 0; Mismatches 379; Indels 12; Gaps 4;
 DB 424 AAGTATCCAAAGCTGTTTAAAGATCGAGGCAATGTGGAAGAGAGGCTGCTGTT 483
 |||||
 DB 596 AAGTATGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 655
 |||||

```

QY 484 CACAGGCTGATCAGCTCTCAAAAGCTGGGAAAGCTGTGTCTCCGGCCATGAAGCTCAAT 543
DB 656 AAGAGGCTTTACCAAGATGCAAGGCTAACCAATCCCTTTCTTCCCTGGACCAATGTCAT 715
QY 544 GATTCTGTACCAAGAGATTTGATTAATTGTACTGTCTCCGAGATTCATTTGGAT 603
DB 716 GACTCCGTACCAAGAGCAAGTTGACAAATCTATAGATGCGGCACTCTCTCCGAT 775
QY 604 GGCCTGAAGAGCAACAGATGTGATGTTGTGTGGGAAACAGTGTGTGTGTGTAT 663
DB 776 GGTCTGATGAGGGCCACTGATGTTATGATGCTGCGCAAGGTTGACAGTTGCTCGGTTAT 835
QY 664 GGTGAGTGGGCAAGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 723
DB 836 GGTGATGTGGAGAGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 895
QY 724 ACCGAATGACCCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 783
DB 896 ACGGATGACCCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 955
QY 784 CTAAATGAAGTCACTCCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 843
DB 956 CTGAGAGATGTTGTCTCAGAGGGGAGATCTTGTGTGTGTGTGTGTGTGTGTGTGT 1015
QY 844 GTGACACGGGAGACCTTGATGTCATGAAAGAGTTGTGTGTGTGTGTGTGTGTGTGT 903
DB 1016 ATCATGCTGACCAACATGAGAGAGATGAGAAAGAAATGCTGTGTGTGTGTGTGTGT 1075
QY 904 TCCACACAGAAATCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
DB 1076 TTTGACAGAGATTTGACATGCTAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1135
QY 961 GTACGTTCTAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1017
DB 1136 ATCAACCCCAAGACCTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1195
QY 1018 GCAGAGGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1074
DB 1196 GCTGAGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1255
QY 1075 ATCAGACGCAACAGAGCTTTGGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1131
DB 1256 TGCTCTTTCACCAACAGGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1315
QY 1132 CGATACAGCAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1191
DB 1316 AAGTATGAGAAAGATTTACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1375
QY 1192 CATCTGCCATATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1251
DB 1376 CACTTGGGCAAGCTCGAGCAAGCTTACAAACCTCAGCCCTTCAGAGGGGATCATC 1435
QY 1252 GGACCTCAACAAAATGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1296
DB 1436 AGCGTCCCATCGAGGGTCTCTTCAAGCCACTCTACAGGTAC 1480

```

RESULT 7

US-08-896-005-2

Sequence 2, Application US/08896005

Patent No. 5854023

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Lal, Preeti

APPLICANT: Shah, Puri

TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

```

COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/896, 005
  FILING DATE: Filed Herewith
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:
    FILING DATE:
ATTORNEY/AGENT INFORMATION:
  NAME: Billings, Lucy J.
  REGISTRATION NUMBER: 36,749
  REFERENCE/DOCKET NUMBER: PF-0337 US
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 415-855-0555
    TELEFAX: 415-845-4166
  INFORMATION FOR SEQ ID NO: 2:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 2226 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    IMMEDIATE SOURCE:
      LIBRARY: BLADTUT04
      CLONE: 1519044
US-08-896-005-2

```

Query Match 9.0%; Score 117; DB 2; Length 2226;
Best Local Similarity 100.0%; Pred. No. 36-27;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1180 GTTGCACGCTTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1239
DB 1 GTTGCACGCTTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY 1240 GCAAAATATCTGGGACTCAACAAAATGGGCAATTCAACTAATTTATTCAGATAC 1296
DB 61 GCAAAATATCTGGGACTCAACAAAATGGGCAATTCAACTAATTTATTCAGATAC 117

```

RESULT 8

US-09-347-878-3

Sequence 3, Application US/09347878C

Patent No. 6376210

GENERAL INFORMATION:

APPLICANT: Yuan, Chong

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES

FILE REFERENCE: 25885-1651

CURRENT APPLICATION NUMBER: US/09/347, 878C

CURRENT FILING DATE: 1999-07-06

NUMBER OF SEQ ID NOS: 75

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 2226

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)..(2226)

OTHER INFORMATION: Polynucleotide encoding human

OTHER INFORMATION: S-adenosyl-5-homocysteine hydrolase (SAHH) derived

PUBLICATION INFORMATION:

PATENT DOCUMENT NUMBER: 08/896, 005

PATENT FILING DATE: 1997-07-17

PUBLICATION DATE: 1998-12-29

US-09-347-878-3

Query Match 9.0%; Score 117; DB 4; Length 2226;

Best Local Similarity 100.0%; Pred. No. 3e-27;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1180 GTTCCAGGCTGATCTGATTCATTTGATTCAGCCACCTTACAGAGGTGACAGATGACCAA 1239
|||||
DB 1 GTTCCAGGCTGATCTGATTCATTTGATTCAGCCACCTTACAGAGGTGACAGATGACCAA 60
|||||
QY 1240 GCAAAATATCTGGAGCTCAACAAAATGGCCATTCAAACCTAATATTATACAGATAC 1296
|||||
DB 61 GCAAAATATCTGGAGCTCAACAAAATGGCCATTCAAACCTAATATTATACAGATAC 117
|||||
RESULT 9
US-08-204-740-8
Sequence 8, Application US/08204740
Patent No. 5753432
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allgretti & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,740
FILING DATE: 04-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5753432nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1234
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ. ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-204-740-8
Query Match 5.6%; Score 72.6; DB 1; Length 289;
Best Local Similarity 58.6%; Pred. No. 1.6e-13;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 32 ACATCAAGCAGCAGAGATTGGACGCCGAGATTGAGATTGACAGCAAGACATGCTG 91
|||||
DB 75 ACATCGGCTGCTGCTGCTGGAGCAGAGCCCTGACATTGCTGAGAAAGAGATGCCG 134
|||||
QY 92 CTCTGATTACACAGAAACGCTCTAGGGGAGAGCCCTTGGCTGCTAAATAG 151
|||||
DB 135 GCGTGAAGCTGATGCGGAGCGGCTGCTGAGCCCTTCAAGCCACTGAAGGCGCCGCTG 194
|||||
QY 152 TGGGCTGTACACATACACAGCCAGCAGCGGTGATTGATGAGACACTGTGCTG 211
|||||
DB 195 CTGCTGCTGCTGACATACAGCCGTGAGACGCGCTGCTCTCATTTGAGACCCCTGCTAC 254
|||||

QY 212 GGGCTCAGTCCGCGCTGCTGCTGCTGATGATCTAC 246
|||||
DB 255 GTGCTGAGGTGAGTGCTCAGCTGCAACATCTTC 289
|||||
RESULT 10
US-09-081-167A-8
Sequence 8, Application US/09081167A
Patent No. 6083745
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive, 32nd Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,167A
FILING DATE: 18-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 6083745nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-KK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ. ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-081-167A-8
Query Match 5.6%; Score 72.6; DB 3; Length 289;
Best Local Similarity 58.6%; Pred. No. 1.6e-13;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 32 ACATCAAGCAGCAGAGATTGGACGCCGAGATTGAGATTGACAGCAAGACATGCTG 91
|||||
DB 75 ACATCGGCTGCTGCTGCTGGAGCAGAGCCCTGACATTGCTGAGAAAGAGATGCCG 134
|||||
QY 92 CTCTGATTACACAGAAACGCTCTAGGGGAGAGCCCTTGGCTGCTAAATAG 151
|||||
DB 135 GCGTGAAGCTGATGCGGAGCGGCTGCTGAGCCCTTCAAGCCACTGAAGGCGCCGCTG 194
|||||
QY 152 TGGGCTGTACACATACACAGCCAGCAGCGGTGATTGATGAGACACTGTGCTG 211
|||||
DB 195 CTGCTGCTGCTGACATACAGCCGTGAGACGCGCTGCTCTCATTTGAGACCCCTGCTAC 254
|||||
QY 212 GGGCTCAGTCCGCGCTGCTGCTGCTGATGATCTAC 246
|||||
DB 255 GTGCTGAGGTGAGTGCTCAGCTGCAACATCTTC 289
|||||

RESULT 14
US-08-204-740-6
Sequence 6, Application US/08204740
Patent No. 5753432
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazon, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allgretti & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,740
FILING DATE: 04-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: NO. 5753432nam, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000

RESULT 15
 US-09-081-167A-6
 : Sequence 6, Application US/09081167A
 : Patent No. 6083745
 :
 : GENERAL INFORMATION:
 : APPLICANT: Gudkov, Andrei
 : APPLICANT: Kazarov, Alexander
 : APPLICANT: Mazo, Ilya
 : APPLICANT: Roninson, Igor B
 : TITLE OF INVENTION: Methods for Identifying Genetic
 : TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
 : TITLE OF INVENTION: Growth in Cancer Cells
 : NUMBER OF SEQUENCES: 15
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
 : STREET: 300 S. Wacker Drive, 32nd Floor
 : CITY: Chicago
 : STATE: Illinois
 : COUNTRY: USA
 : ZIP: 60606
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/081,167A
 : FILING DATE: 18-MAY-1998
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: No. 6083745naa, Kevin E
 : REGISTRATION NUMBER: 35,303
 : REFERENCE/DOCKET NUMBER: 93,354-KK
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 312-913-0001
 : TELEFAX: 312-913-0002
 :
 : TELEX:
 : INFORMATION FOR SEQ ID NO: 6:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 285 base pairs
 : TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-081-167A-6

Query Match 5.3%; Score 68.4; DB 3; Length 285;
Best Local Similarity 56.8%; Pred. No. 3.5e-12;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 25 GTGAGAACATCAAGCAGAGAGATTGAGCGCGGAGATTGAGATTGACAGACAGAC 84
DB 64 GTGCGGACATCGGACTGCGCGCTGGGGACGAGAGCTCTGATATAGCTGAGAGATGAG 123
QY 85 ATGCTGCTCTGATTTCACACAGAAACGTCTCAGGGGAGAGAGCCCTTGGCTGTGCT 144
DB 124 ATGCCAGGGTTGATGCGCATGCGGAGATGTACTCAGCCTCCAGCCACTGAAGGGTGT 183
QY 145 AAAATAGTGGCTGTACACATCACAGCCACAGAGCGGTGTGATTGAGACACTCTGT 204
DB 184 CGCATTGCTGGCTGCTGCGCATGACCGGTGAGAGCTGTCTTCATTGAGACTCTGTG 243
QY 205 GCCCTGGGGGCTCAGTGCCTGCTGCTGTGCTGTGATACATCTAC 246
DB 244 GCCCTGGGTGCTGAGGCGCGGTGCTCAGCTGCACATCTTC 285

Search completed: April 21, 2003, 04:34:47
Job time : 2035.93 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: April 21, 2003, 02:39:45 ; Search time 259.419 Seconds

(without alignments)
5018.577 Million cell updates/sec

Title: US-09-782-051-1_COPY_549_1844

Perfect score: 1296

Sequence: 1 tccaaggcagcagcagcattt.....aaccaattatcacgataac 1296

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1296	100.0	2563	10 US-09-782-051-1	Sequence 1, Appl1
2	461.4	35.6	553	10 US-09-867-701-2023	Sequence 2023, Ap
3	406.4	31.4	636	10 US-09-879-536-850	Sequence 850, App
4	402.4	31.0	2300	10 US-09-925-301-217	Sequence 217, App
5	402.4	31.0	2429	10 US-10-044-090-344	Sequence 344, App
6	337	26.0	384	10 US-09-880-107-346	Sequence 346, App
7	241.6	18.6	1461	10 US-09-759-990-1	Sequence 1, Appl1
8	225.6	17.4	1422	10 US-09-738-626-836	Sequence 836, Appl1
9	225.6	17.4	1557	10 US-09-746-660A-97	Sequence 97, Appl1
10	190	14.7	1877	12 US-10-021-121-1	Sequence 1, Appl1
11	186	14.4	1396	9 US-09-746-660A-101	Sequence 101, App
12	174	13.4	720	9 US-09-738-626-838	Sequence 838, App
13	172.4	13.3	3830	9 US-10-037-598-26	Sequence 26, Appl1
14	172.4	13.3	513509	9 US-09-754-853A-4	Sequence 4, Appl1
15	168.2	13.0	433	10 US-09-960-352-4421	Sequence 4421, Ap
16	158	12.2	721	10 US-09-925-300-439	Sequence 439, App
17	150.8	11.6	382	10 US-09-960-352-11499	Sequence 11499, A
18	138.4	10.7	284	10 US-09-920-300A-1423	Sequence 1423, Ap
19	138.4	10.7	284	12 US-10-033-528-1423	Sequence 1423, Ap

20	137.6	10.6	424	10 US-09-960-352-9160	Sequence 9160, Ap
21	136.8	10.6	288	9 US-09-736-457-1179	Sequence 1179, Ap
22	136.8	10.6	288	9 US-09-902-941-1179	Sequence 1179, Ap
23	136.8	10.6	288	9 US-09-849-626-1179	Sequence 1179, Ap
24	136.8	10.6	288	9 US-10-017-754-1179	Sequence 1179, Ap
25	136.4	10.5	288	10 US-09-998-598-2191	Sequence 2191, Ap
26	136.2	9.9	352	10 US-09-960-352-12928	Sequence 12928, A
27	124.4	9.6	283	10 US-09-960-352-7006	Sequence 7006, Ap
28	114.4	8.8	793	10 US-09-966-881-9	Sequence 9, Appl1
29	111.4	8.6	708	9 US-09-738-626-837	Sequence 837, Appl1
30	108.2	8.3	389	10 US-09-960-352-10073	Sequence 10073, A
31	102.4	7.9	728	10 US-09-770-149-117	Sequence 117, App
32	100.6	7.8	237	10 US-09-960-352-1145	Sequence 1145, A
33	98.4	7.6	263	10 US-09-923-876-477	Sequence 477, App
34	96	7.4	410	10 US-09-960-352-1365	Sequence 1365, Ap
35	94.2	7.3	253	10 US-09-960-352-7284	Sequence 7284, Ap
36	91.4	7.1	399	10 US-09-878-574-838	Sequence 838, App
37	81.4	6.3	219	10 US-09-960-352-12043	Sequence 12043, A
38	79.2	6.1	456	10 US-09-960-352-2551	Sequence 2551, Ap
39	78.2	6.0	835	10 US-09-770-445-679	Sequence 679, App
40	75.2	5.8	405	10 US-09-960-352-5161	Sequence 5161, Ap
41	72.6	5.6	289	10 US-09-799-946-8	Sequence 8, Appl1
42	69.8	5.4	347	10 US-09-960-352-2035	Sequence 2035, Ap
43	68.4	5.3	285	10 US-09-799-946-6	Sequence 6, Appl1
44	67.6	5.2	163	10 US-09-878-574-9524	Sequence 9524, A
45	65.8	5.1	428	10 US-09-960-352-10555	Sequence 10555, A

ALIGNMENTS

RESULT 1
US-09-782-051-1
; Sequence 1, Application US/09782051
; Patent No. US20020035078A1
; GENERAL INFORMATION:
; APPLICANT: Hart, Derek N J
; TITLE OF INVENTION: Enzyme having S-adenosyl-L-homocysteine hydrolase
; FILE REFERENCE: 24305 MRB
; CURRENT APPLICATION NUMBER: US/09/782,051
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: PCT/NZ97/00133
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: NZ 299507
; PRIOR FILING DATE: 1996-10-04
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2563
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(1847)
; OTHER INFORMATION: Open reading frame extends without a stop codon
; OTHER INFORMATION: for the full 5' nucleotide sequence. The
; OTHER INFORMATION: Initiation codon has yet to be identified.
US-09-782-051-1

Query Match 100.0%; Score 1296; DB 10; Length 2563;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TCCAAGGCGAGCAGCAATTTCTGTGTAACATCAAGCAGCAGCAATTTGACGCCGG	60
DB	549	TCCAAGGCGAGCAGCAATTTCTGTGTAACATCAAGCAGCAGCAATTTGACGCCGG	608
QY	61	GAGATTGAGTTGAGAGCAGCATGTCTGCTGATTCACATCAGAAAGTGTCTAG	120
DB	609	GAGATTGAGTTGAGAGCAGCATGTCTGCTGATTCACATCAGAAAGTGTCTAG	668
QY	121	GGGAGAGACGCTTGCTGTGCTAAATGATGGGCTGTACACATCAGCCACAGCA	180

D	669	GGGAGAGAGCCCTGGCTGGTGTCTAAATAATAGGGCTGTACACATATCACAGCCACACA	728
Q	181	GGGGTGTGATTAGACACTCTGTGCTTGGGCGCTCACTGCCGCTGGTCTGTGTATAC	240
D	729	GGCGTGTGATTAGACACTCTGTGCTTGGGCGCTCACTGCCGCTGGTCTGTGTATAC	788
Q	241	ATCTACTCACTAGAGATGAAGTGTGAGACTGGAGACTGGCGGTGGAGTTGCAGTGTTC	300
D	789	ATCTACTCACTAGAGATGAAGTGTGAGACTGGAGACTGGCGGTGGAGTTGCAGTGTTC	848
Q	301	GCTTGAAGGCGAGTCCAGAAATGACTCTGTGTGTATTGACCGCTGTGTGAACATG	360
D	849	GCTTGAAGGCGAGTCCAGAAATGACTCTGTGTGTATTGACCGCTGTGTGAACATG	908
Q	361	GATGGGTGGCAGGCGCAACATGATCTCTGGATGATGGGGGAGACTTAAACCACATGGGTTAT	420
D	909	GATGGGTGGCAGGCGCAACATGATCTCTGGATGATGGGGGAGACTTAAACCACATGGGTTAT	968
Q	421	AGAAGATCCAAACGTTTAAAGATCCAGGCAATGTGGAAGAGCGTCACTGT	480
D	969	AGAAGATCCAAACGTTTAAAGATCCAGGCAATGTGGAAGAGCGTCACTGT	1028
Q	481	GTCACAGCGCTATCACTCTCCAAAGCTGGGAGCTGTGTTCGGCCATGAACCTC	540
D	1029	GTCACAGCGCTATCACTCTCCAAAGCTGGGAGCTGTGTTCGGCCATGAACCTC	1088
Q	541	AATGATCTGTATCCAAACAGAGTTTGATTAACCTGTACTGCTGCCGAGAAATCCATTGTG	600
D	1089	AATGATCTGTATCCAAACAGAGTTTGATTAACCTGTACTGCTGCCGAGAAATCCATTGTG	1148
Q	601	GATGCGCTGAAGAGAGACACAGATGTGATGTTTGTGGGAAACAAGTGTGTGTGGC	660
D	1149	GATGCGCTGAAGAGAGACACAGATGTGATGTTTGTGGGAAACAAGTGTGTGTGGC	1208
Q	661	TATGTGAGGTAGGCAAGGCGCTGCTGTGCTGTCCAAAGCTTGGAGCAATTGTCTAC	720
D	1209	TATGTGAGGTAGGCAAGGCGCTGCTGTGCTGTCCAAAGCTTGGAGCAATTGTCTAC	1268
Q	721	ATTACCGAAATCGAACCCCATCTGTGCTGTGCAGAGCGCTGCATGGATGGGTTCAAGGTTGTA	780
D	1269	ATTACCGAAATCGAACCCCATCTGTGCTGTGCAGAGCGCTGCATGGATGGGTTGTA	1328
Q	781	AAGTAATATAAGCATCCGGGCAAGTCCGATGTCCTTAATCTTGACAGAAATTAAGAT	840
D	1329	AAGTAATATAAGCATCCGGGCAAGTCCGATGTCCTTAATCTTGACAGAAATTAAGAT	1388
Q	841	GTAATGACAGGGGAGCACTTGTGATCCCATGAAGAAACAGTTGTATGCAATATGGC	900
D	1389	GTAATGACAGGGGAGCACTTGTGATCCCATGAAGAAACAGTTGTATGCAATATGGC	1448
Q	901	CACCTCAACACAGAAATCATGTGTGACACGCTCCGACTCCGGAGCTGACGTGGAGCGA	960
D	1449	CACCTCAACACAGAAATCATGTGTGACACGCTCCGACTCCGGAGCTGACGTGGAGCGA	1508
Q	961	GTAATCTCAGGAGGAGCATGTGATCTGGCCAGATGGCAAGAGAGTTGTCTCTCTGGCA	1020
D	1509	GTAATCTCAGGAGGAGCATGTGATCTGGCCAGATGGCAAGAGAGTTGTCTCTCTGGCA	1568
Q	1021	GAGGGTGTCTACTCAATTTGAGCTGCTCACAGTTCCACACTTGTCTGTCTCATACACA	1080
D	1569	GAGGGTGTCTACTCAATTTGAGCTGCTCACAGTTCCACACTTGTCTGTCTCATACACA	1628
Q	1081	GGCAACAACAGGCTTGGCAGATGAAGAACTTAATAGCAACCGGAGGGCGGATACAG	1140
D	1629	GGCAACAACAGGCTTGGCAGATGAAGAACTTAATAGCAACCGGAGGGCGGATACAG	1688
Q	1141	CAGAGTGTACTGTCTCTTAAGAAATGATGATACGTTGCCAGGTGTGCATGTGCA	1200
D	1689	CAGAGTGTACTGTCTCTTAAGAAATGATGATACGTTGCCAGGTGTGCATGTGCA	1748
Q	1201	TCAATTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAAC	1260

Db 1749 TCATTTCATGCCACCTTACAGAGCGTACAGATGACCAAGCAAAATATCTGGACTCAC 1808

Qy 1261 AAAATGCGCATTCAAACCTAATTATTACAGATAC 1296

Db 1809 AAAATGCGCATTCAAACCTAATTATTACAGATAC 1844

```
US-09-867-701-2023
RESULT 2
Sequence 2023: Application US/09867701
Patent No. US20020132237A1
GENERAL INFORMATION:
APPLICANT: Agiate, Paul A.
APPLICANT: Jones, Robert
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.497
CURRENT APPLICATION NUMBER: US/09/867,701
CURRENT FILING DATE: 2001-05-29
NUMBER OF SEQ ID NOS: 10912
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2023
LENGTH: 553
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1) --(553)
OTHER INFORMATION: n = A,T,C or G
US-09-867-701-2023
```

Query Match	35.68;	Score 461.4;	DB 10;	Length 553;
Best Local Similarity	97.38;	Pred. No. 8e-147;		
Matches 510;	Conservative 0;	Mismatches 9;	Indels 5;	Gaps 4

OY	774	GGTGTAAAGCTAAATTAATCATCCGGCAAGTCGATGCTGTAATAACTGGACAGGAAA	833
Db	1	GGTGTAAAGCTAAATTAATCATCCGGCAAGTCGATGCTGTAATAACTGGACAGGAAA	60
OY	834	TAAAGATGTAGTCAGCAGGAGACCTTGGATGGCATGAAAAACAGTTGTATCGATGCAA	893
Db	61	TAAAGATGTAGTCAGCAGGAGACCTTGGATGGCATGAAAAACAGTTGTATCGATGCAA	120
OY	894	TATGGGCACTCCAAACACAGAAATTCATGTGACACACCTCCGACCTCGGAGCTACGTG	953
Db	121	TATGGGCACTCCAAACACAGAAATTCATGTGACACACCTCCGACCTCGGAGCTACGTG	180
OY	954	GGAGCGAGTACGTTCTCAGGTGAGCAACATGTCATCTGGCCAGATGGCAACGAGTTGTCT	1013
Db	181	GGAGCGAGTACGTTCTCAGGTGAGCAACATGTCATCTGGCCAGAT - GNAACAGAGTTGTCT	239
OY	1014	CTTGCGAGAGGGTGCTGCTACTCAATTGAGCTGCTCACAAGTCCACACTTGTGTCTG	1073
Db	240	CTTGCGAGAGGGTGCTGCTACTCAATTGAGCTGCTCACAAGTCCACACTTGTGTCTG	299
OY	1074	CATCACAGGCCACAACACAGGCTTTGGGACTGATAGACTCTATAATGCACCCGAGGGCG	1133
Db	300	CATCACAGGCCACAACACAGGCTTTGGGACTGATAGAACTCTATAATGCACCCGAGGGCG	359
OY	1134	ATAACAGCAGATGCTGACTGTCTCCCTTAAGAAA - TGGATATAATGCTTGGCAGCTGC	1192
Db	360	ATAACAGCAGATGCTGACTGTCTCCCTTAAGAAAATGGATGAATACGTGGCAGCTGC	419
OY	1193	ATTCGCATCATTTTGATGATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	1252
Db	420	ATTCGCATCATTTTGATGATGCCACCTTACAGAGCTGACAGATGACCAAGC - AAAATATCTGG	478
OY	1253	GACTCACAACAAAATGGGCACTTCAAACTTAATTAATACAGATAC	1296
Db	479	GACTCACAACAAA - TGGGCACTTCAAACTTAATTAATTAATACAGATAC	520

```

RESULT 3
US-09-879-536-850/c
; Sequence 850, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Asile, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Carino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 850
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(636)
; OTHER INFORMATION: n = A,T,C or G
US-09-879-536-850

Query Match          31.4%, Score 406.4; DB 10; Length 636;
Best Local Similarity 99.5%; Pred. No. 5.2e-128;
Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 159 TACACATCATCAGCCAGACAGC-GGTGTGATGAGACACTGTGCGCCGGGGGCGTC 217
DB 420 TACACATCATCAGCCAGACAGCAGCGGGTGTGATGAGACACTGTGCGGGGGCTC 361
QY 218 AGTGGCGGTGCTGTCTGTAACTACTCACTCAGATGAATAGCTGACGACTGG 277
DB 360 AGTGGCGGTGCTGTCTGTAACTACTCACTCAGATGAATAGCTGACGACTGG 301
QY 278 CTGAGCGTGAAGTTCAGCTGTTCGCTTGAAGGCGAGTCAAGAATGACTTCTGGTGT 337
DB 300 CTGAGCGTGAAGTTCAGCTGTTCGCTTGAAGGCGAGTCAAGAATGACTTCTGGTGT 241
QY 338 GTATGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCCTGGATGATGGGG 397
DB 240 GTATGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATCCTGGATGATGGGG 181
QY 398 GAGACTTAACCCACTGGGTTTATAGAAGTATCCAACGCTGTTTAAAGAATCCGAGCA 457
DB 180 GAGACTTAACCCACTGGGTTTATAGAAGTATCCAACGCTGTTTAAAGAATCCGAGCA 121
QY 458 TTGCGAAGAGAGCGTGAAGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGC 517
DB 120 TTGCGAAGAGAGCGTGAAGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAGC 61
QY 518 TCTGTGTCCGGCCATGAACGTCATGATCTGTATACCAAAAGAAGTTGATTAATCTGT 577
DB 60 TCTGTGTCCGGCCATGAACGTCATGATCTGTATACCAAAAGAAGTTGATTAATCTGT 1

```

```

GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 217
LENGTH: 2200
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2188)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-217

Query Match          31.0%; Score 402.4; DB 10; Length 2200;
Best Local Similarity 58.4%; Pred. No. 2.6e-126;
Matches 741; Conservative 0; Mismatches 521; Indels 6; Gaps 2

QY 32 ACATCAAGCAGGACAGAAATTTGAGACGCCGGAGATTGAGATTGACAGACGACATGCTG 91
DB 108 ACATGAGGCGCTGCTGCTCGGGAGACGACGAGCCCTGAGACTTCTGAGACGAGATGCCG 167
QY 92 CTCGTATTTCACTCAGGAACGCTGCTCAGGGGAGAGAGCCCTTGGCTGCTTAATATG 151
DB 168 GCCTGATGCGTATGCGGGGAGCGGTACTCGGCTCCAGGACACTGGAAGGCGCCGCGATG 227
QY 152 TGGGCTGTAACAACAATCACAGCCGACAGACGCGGCTGTGATTGAGACACTCTGGCCCTG 211
DB 228 CTGGCTGCGCTGACATGACACCGTGAGAGCGCCGCTCCTCTTATGAGACCTCGTACCCCTG 287
QY 212 GGGCTCAGTGCACCGCTGCTGCTGTTTAACTATCTACTCAACTCAGATGAAATGAGTGCAG 271
DB 288 GTGCTGAGGTGACAGTGTCCAGCTGCACATCTTCTCCACCGACGACATGCGGGCGCTG 347
QY 212 CACTGCTGAGGCTGGAAGTTCGAGTTCGCTTGGAAGGGCGAGTCACAGATGACTCT 331
DB 348 CCATTGCCAAGGCTGGCACTTCGCGGTATGCTCGGAAGGGCGCAAAACGAGCGAGATACC 407
QY 332 GGTGCTGATATGACCGCTGTGTGAACATGAGAGGGTGGCGGACCAACATGATCCTGGATG 391
DB 408 TGTGTGCACTTGAGACAGACCCGTGTACTTAAAGAGCGGCGCCCTCAACATGATTCGGAGG 467
QY 392 ATGGGAGAGACTTAAACCACTGGTATTATAAAGATATCAACAGCTTTAAGAAGATCC 451
DB 468 ACGGGGCGGACCTCACCAACTCATTCCACACCAAGTACCCGAGAGCTCTGCCAGGATCC 527
QY 452 GAGGCAATTGTGAGAGAGAGCGTGAAGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTG 511
DB 528 GAGGCACTCTGAGAGAGACACGACTGGGGGTCCACAAACCTTAAGAGATGAGGCCAAAG 587
QY 512 GGAACCTCTGTGTTCGGGCGACGTAACGTAAAGATCTGTTACCAACAGATGTTGANA 571
DB 588 GGATCTCAAGGTGCTGCTGCACATCAATGTCAAAAGACTCCGTCACCAAGAGAGTTTGACA 647
QY 572 ACTTGTAAGCTGCGCGAGAAATCCATTTTGGATGGGCTTAAGAGAGACACAGATGTGATGT 631
DB 648 ACCTGTATAGGCTGCGCGGAGATCCCTCATAGATGAGCATCAAGGGGCGCACAGATGTGATA 707
QY 632 TTGTTGGGGAACAAGTGTGTGTGTGCTATAGTGTAGGTAGGCAAGGGCTGCTGTGCTG 691
DB 708 TTGCGCGCAAGGTAGCGGTGAGACAGGCTATGTGATGTGGCAAAAGGCTGTGCCAAG 767
QY 632 CTCCTCAAGGCTTTGGAGCAATTGTCTCATTAACGAAATGCAACCCCATCTGTGCTGTC 751
DB 768 CCTCTGGGGGTTTTCGAGGCCCGCGTATATATACCGAGATTGACCCCATCAACGCACTGC 827

```

QY 752 AGGCGTGGATGGATGGGCTCAGGCGGTAAAGCTAATGAATGATCCGGCAAGTCGATG 811
 Db 828 AGGCTGCCATGAGGCGCTATGAGTGTACCAACCATGGATGAGGCTGTGCAGAGGGGCAACA 887
 QY 812 TCGTATTAATCTTGCACAGGAATAATGAATGTAGTGACACGGGAGCACTTGGATCGCATGA 871
 Db 888 TCTTTGTACCAACCAAGGCTGTATGACATCATCTCTGGCCGGCACTTTGAGACAGATG 947
 QY 872 AAACAGTTGTATCTGATTCATATGCGCCACTCCAAACAGAAATGAGTGTACCAAGCC 931
 Db 948 AGGATGATGCGCATTTGTGTATTAACATTTGACACTTTGACGTGGAGATCGATGTCAAGTGGC 1007
 QY 932 TCCGCACTCCGAGCGAGCTGGGAGCGAGTACGTTCTCAGGTGGACCATGTCACTGTGCG 991
 Db 1008 TCAAGAGAGAGCCCGGTGGAGAGGTGAACATCAAGCCGAGGTGGAGCCGGATGCGTGA 1067
 QY 992 CAGATGCCAAGAGATGTGTCTCTCTGCGCAGAGGGGTGTCTACTCAATTTGAGCTGTCCA 1051
 Db 1068 AGAATGGGGCGCGCATCACTCTGCTGCGAGGGGTGGCGGTGTCMACTGGTGTGTGCCA 1127
 QY 1052 CAG---TTCCCACTCTGTCTGTCTGTCATCATCAGCCACACACAGGCTTTGGACATGATG 1108
 Db 1128 TGGGCCACCCCACTCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1187
 QY 1109 AACTATATATGACACCCGAGGGCGATACAGAGAGATGTGTACTTCTCTTCAAGAAA 1168
 Db 1188 AGCTGTGGACCCATCCAGACAG---TACCCCGTTGGGGTTCAATTTCTGCCCCAAGAGC 1244
 QY 1169 TGGATATATACGTGTGCGACGTTGCTGATCTGCTGATCTGATGATGATGATGATGATG 1228
 Db 1245 TGGATAGGCGAGTGTGTAAGCCCACTGGGCAAGCTGTAATGTAGTGTGACCAAGCTTAA 1304
 QY 1229 CAGATGACCAAGCAAAATATCTGGGACTCAACAAAATGGGGCACTTCAAACTATATAT 1288
 Db 1305 CTGAGAGAGCAAGCCAGTACCTGTGGCATGTCTGTATGCGCCCTTCAAGCCGATCACT 1364
 QY 1289 ACAGATAC 1296
 Db 1365 ACCGCTAC 1372
 ~
 RESULT 5
 US-10-044-090-344
 ; Sequence 344, Application US/10044090
 ; Patent No. US20020137081A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Olga Bandman
 ; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
 ; FILE REFERENCE: PA-0028 US
 ; CURRENT APPLICATION NUMBER: US/10/044, 090
 ; CURRENT FILING DATE: 2002-01-09
 ; NUMBER OF SEQ ID NOS: 850
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 344
 ; LENGTH: 2429
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc. feature
 ; OTHER INFORMATION: Incyte ID No. US20020137081A1 1468237CH1
 US-10-044-090-344
 ~
 Query Match 31.0%; Score 402.4; DB 12; Length 2429;
 Best Local Similarity 58.4%; Pred. No. 2.8e-126;
 Matches 741; Conservative 0; Mismatches 521; Indels 6; Gaps 2;
 QY 32 ACATCAAGCAGGAGAAATTTGGACGCGGAGATTGAGATTGCAGAGCAAGCATGTCTG 91
 Db 108 ACATGCGCGCTGCTGCTGGGAGCGCAAGGCCCTGACATGCTGAGAAACGAGATGCCGG 167
 QY 92 CTGTGATTACATCAGAAACGTCTCAGGGGGAGAAACCCCTTGGCTGTGTAAATG 151

Db	168	GCCTGATGCGGATACCGGGAGCGGCTACTCGACCTCCAAAGCCATGTGAAGGGCGCCCGCATCG	227
Oy	152	TGGCGTGTACACACATATACAGAGCCCAACAGACGGGTGTGATTATGACACTCTGTGCGCTGG	211
Db	228	CTGGCTGCTGTGCACATATACCGGTGAGAGCGGCCCTCTCTCATTTGAGACCCCTGTCACTCCGG	287
Oy	212	GGGCGTACGTGCGCGTGTGCTGTGTATATCAATCTACTCACTCAGATGTGAATGATGTGAG	271
Db	288	GTGCTGAGGTGCACTGGTGTCCAGCTGCACATCTTCTTCCACCCAGAACCATGTGGCGGGCTG	347
Oy	272	CACGTGCTGAGAGCTGGAGTTGCAAGTGTTCGTTCTGGAAAGGGCGAGTCAGAAGATGACTCT	331
Db	348	CCATTGCCAAGGCGTGGGATTCGCGGTGTATCCTGTGAAGGGCGAAAGCGAGAGATGAC	407
Oy	332	GGTGTGTATTGACCGCTGTGTGAACTGATGGGTGGCAGGCCAACATGATCCGGATG	391
Db	408	TGTGTGTGATTGACACACACCTGTACTTCAAGAGCGGGCCCCCTCAACATGATTTGTGGAG	467
Oy	392	ATGGGAGGACTTAAACCCACTGGGTTTATTAAGAGTATCCAAACGTGTTTAAGAAAGATCC	451
Db	468	AAGGGGGGAGACTCAACCACTCATCCACACAAAGTACCGCGAGCTTCTGCGAGGATCTC	527
Oy	452	GAGCAATTGTGAAAGAGCGCTAGCTGTTCACAGCTGTATCAAGCTTCCAAAGCTG	511
Db	528	GAGGCACTCTGAGAGAGACCCAGCTGGGGTCCACAACTCTCAAGAGATGGCCAAAG	587
Oy	512	GGAACCTGTGTTCGCGCCATGAAAGCTCAATGTTCTGTACCAACAGAAATTGATA	571
Db	588	GGATCCTTAAGGTGCTGTCCATCATGTGCATGCTCTGTCACCAAGAGACAGATTGACA	647
Oy	572	ACTTGTACTGCTGCGAGAAATCATTTTGGATGAGCCCTGAAGAGAGACACAGATGTATGT	631
Db	648	ACCTGTATAGGCTGGCGGGAGTCCCTCATATAGTGCATCAAGGGGCGACAGATGTATGA	707
Oy	632	TTGTGGGAAACAGATGTTGTGTGTGTGCTATGTGTAGGTAGGCAAGGGCTGTGTGCTG	691
Db	708	TTGCGCGCAAGATGACGGTGTGTATACAGGCTATGTGTATGTGGCAAGGGCTGTGGCCAG	767
Oy	692	CTCTCAAAAGCTTTGGAGCAATTTGTCATCATTTACGAAATCGAACCCATCTGTGCTCG	751
Db	768	CCCTCGGGGTTTGTGGAGAGCCCGGCTCATCATCACCGAGATTGACCCCATCAACGCACTGC	827
Oy	752	AGCGCTGCATGATGGTGGTTCAGGGTGTGTAAGCTAAATGAAGTATCCGGCAAGTCAGT	811
Db	828	AGGCTGCAATGAGAGGGCTATAGAGTACACACATGATGAGGCTGTGACAGAGGCAACA	887
Oy	812	TGCTAATACTTGCACAGAAATTAAGATGTATGAGACAGGGAGCACTTGGATCGATGA	871
Db	888	TCTTTGTACACCAACAGAGCTGTATTGACATCATCTTTGGCCGGCACTTTTGACAGATGA	947
Oy	872	AAACAGTTGATGATGACAAATATGGGCACTCCACACAGAAAGATGATGACAGGCC	931
Db	948	AGGATGATGCCCTTGTGTGTATGATGACATTTGACCTTGTACGTGGAAATGATGTCAATGGC	1007
Oy	932	TCCGCACTCCGAGCTGACGTGGGAGCAGATACGTTTTCAGGTGGACATGTCATCTGGC	991
Db	1008	TCAAGAGAACGCCCTGAGAAAGTGAATCATCAAGCCGAGAGTGGACCGGTATCCGGTTGA	1067
Oy	992	CAGATGCAACAGATGTCTCTCTGSCAGAGGTCGTACTCATTTGAGTGAAGTCTCCA	1057
Db	1068	AGCAATGGGGCGCATCATCTGCTGGCGCGAGGGTCCGGCTGATCAACCTGGTTGTGCCA	1127
Oy	1052	CAG---TTCCCACTTTTGTCTGTCTCATCAAGCCACAAACAGCGTTTGGCACTATAG	1107
Db	1128	TGGGCGCACCCGCTTGTGTGTATGATTAATCTTTCCACCAACAGATGATGGCCAGATGCA	1187
Oy	1109	AACCTATATATGACCCGAGGGGCGCATCAAGACAGAGATGTGTACTTCTCTTAAGAAA	1167
Db	1188	AGGTGTGGACCCATCAAGCAAG---TACCCCGTTGGGGTTCAATTTCTCTGCCAACAGAGC	1247
Oy	1169	TGATTAATATCTTGCACAGCTGTGATCTGCATCATTTGATGCCCACTTACAGAGCTGA	1227
Db	1245	TGATATAGCAATGGCTGAGAGCCACCTGTGGGCAAGCTGAATGTGAATGTACCAAGCTAA	1307

QY 1229 CAGATGACCAAGCAAAATATCTGGAGACGACGACCAAAATGGCCATTCACCAATATTATT 1288
DB 1305 CTGAGAGCAAGCAAGCCCACTGCTGGCATCTCTGTGATGGCCCTTCAAGCCGATCACT 1364
QY 1289 ACAGATAC 1296
DB 1365 ACCGCTAC 1372

RESULT 6

US-09-880-107-346
; Sequence 346, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 346
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA157401
US-09-880-107-346

Query Match 26.0%; Score 337; DB 10; Length 384;
Best Local Similarity 99.0%; Pred. No. 1.9e-104;

Matches 381; Conservative 0; Mismatches 0; Indels 4; Gaps 4;

QY 444 GAAGATCCAGCATGTGTGGAAGAGAGCGTGAAGTGTTCACAGGCTGTATCAGCTCTC 503
DB 1 GAAGATCCAGCATGTGTGGAAGAGAGCGTGAAGTGTTCACAGGCTGTATCAGCTCTC 60
QY 504 CAAAGCTGGAAAGCTGTGTTCGCGCATGAACGTCATGATTCGTTACCAAGAGAA 563
DB 61 CAAAGCTGGAAAGCTGTGTTCGCGCATGAACGTCATGATTCGTTACCAAGAGAA 119
QY 564 GTTGTATTAATCTGTCTGCTGCGGAGAAATCATTGATGGCTGAAAGAGACACAGA 623
DB 120 GTTGTATTAATCTGTCTGCTGCGGAGAAATCATTGATGGCTGAAAGAGACACAGA 179
QY 624 TGTGATGTTGGGGAAGCAAGTGTGTGTGCTATGATGAGGTAGCAAGGCGTG 683
DB 180 TGTGATGTTGGGGAAGCAAGTGTGTGTGCTATGATGAGGTAGCAAGGCGTG 229
QY 684 -CTGTGCTCTCTC-AAAGCTTTGAGCAATGTCTACATTACCGAAATCGACCCATC 741
DB 240 CCGTGTCTCTCTCAAAAGCTTTGAGCAATGTCTACATTACCGAAATCGACCCATC 299
QY 742 TGTGCTCTCTCAGGCTGCAATGATGAGGTTCAGGCTGT-AAAAGTAAATGATGATCGG 800
DB 300 TGTGCTCTCTCAGGCTGCAATGATGAGGTTCAGGCTGTAAAGTAAATGATGATCGG 359
QY 801 GCAAGTCATGTCGTAATAACTGC 825
DB 360 GCAAGTCATGTCGTAATAACTGC 384

RESULT 7
US-09-759-990-1
; Sequence 1, Application US/09759990

; Patent No. US20020119491A1
; GENERAL INFORMATION:
; APPLICANT: Anticancer, Inc.
; APPLICANT: Xu, Mingxu
; APPLICANT: Han, Qinghong
; TITLE OF INVENTION: HIGH EXPRESSION AND PRODUCTION OF HIGN
; TITLE OF INVENTION: SPECIFIC ACTIVITY RECOMBINANT S-ADENOSYLHOMOCYSTANINE
; TITLE OF INVENTION: (SAHH) AND IMPROVED ASSAYS FOR S-ADENOSYLHOMOCYSTANINE (SAM)
; FILE REFERENCE: 31276-20026.00
; CURRENT APPLICATION NUMBER: US/09/759,990
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/176,444
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1461
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence encoding SAHH
US-09-759-990-1

Query Match 18.6%; Score 241.6; DB 10; Length 1461;
Best Local Similarity 57.8%; Pred. No. 1.6e-71;

Matches 490; Conservative 0; Mismatches 349; Indels 9; Gaps 3;

QY 454 GGCATTGTGGAAGAGACCGTACTGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTGG 513
DB 613 GGTGTTCCGAAAGAGCAACAGAGGTGTCCACCGCTCTACACGCTGAGAAAGAGGCG 672
QY 514 AAGCTCTGTTCGCGGCAAGAGCAATGATTCGTTACCAAGAGTGTGATTAAC 573
DB 673 AAGCTCTGTTCGCGGCAAGAGCAATGATTCGTTACCAAGAGTGTGATTAAC 732
QY 574 TTGTACTGCTGCGGAGATTCATTTTGGATGGCTGAAGAGACACAGATGTGATTT 633
DB 733 ATCTAGGCTGCGCGCCACTCCCTTATGATGATGATATCAACGCTGCTCCATGATATC 792
QY 634 GTTGGGAAGCAAGT 693
DB 793 GGTGGGAAGCAAGT 852
QY 694 CTCGAAGCTTTGAGCAATGT 753
DB 853 CTCGAAGCTTTGAGCAATGT 912
QY 754 GCTGTCATGATGAGGTTCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 813
DB 913 GCTGTCATGATGAGGTTCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 972
QY 814 GTAATACTTGCACAGAGAAATGAAGATGTAGTACAGGAGCACTTGATTCGATGAA 873
DB 973 TTGTTATCATGACAGAGAAATGTAGTATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1032
QY 874 AATAGTTGT 933
DB 1033 GATTAAGCTTTGCTGATCATCTGCGCACTGCGCACTTGTATCAAAATTTGATTAAGTGTG 1092
QY 934 CGCA--CTCGGAGCTGAGAGT 990
DB 1093 ATGAATAATCCAGCATCAAGCAATCCCAATCAAGCAATCAAGCAATCAAGCAATCAAG 1152
QY 991 CCAGATGCAAGAGAGTGTCT 1050
DB 1153 CCAGATGCAAGAGTGTCT 1212
QY 1051 AAGATT--CCCACTTTGTTGT 1107
DB 1213 ACAGGTTCACCATCTTCTGTATGTCAATGTCAATGTCAATGTCAATGTCAATGTCAATGT 1272
QY 1108 GAAGTCTATAATGACACCGAGGCGGATTAAGAGAGATGTGTGTGTGTGTGTGTGTGTGT 1167

DB 1273 GACCTTACGA---AAGAGAGGAAATCTCGAGAGAGGTTTACACACTCCGAAACAT 1329
QY 1168 ATGGATGAATAGCTGGACCTGATCTGCATCTGCAATTTATGCGCCCTACAGACTG 1227
DB 1330 CTGATGAAGAAAGCTGGCTCCCTCCACCTCGATCTCTCATCTCCACCTTACAAACCT 1389
QY 1228 ACAGATGACCAACCAAAATATCTGGAGCTCAACAAAATGGCCATTCAAACCTAATAT 1287
DB 1390 ACACAGAGAGAGGCTGACTACATCAACGTTCCAGTTAGAGGCTCTTCAAAAGTCTGATGCT 1449
QY 1288 TACAGATA 1295
DB 1450 TACCGTTA 1457
RESULT 8
US-09-738-626-836
; Sequence 836, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 836
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: *Corynebacterium glutamicum*
US-09-738-626-836
Query Match 17.4%; Score 225.6; DB 9; Length 1422;
Best Local Similarity 56.8%; Pred. No. 4.6e-66;
Matches 477; Conservative 0; Mismatches 354; Indels 9; Gaps 3;
QY 463 GAAGAGAGCGTGAAGTGTTCACAGGCTGTATGAGCTCTTCAAGCTGAGGAGCTGTGT 522
DB 586 GAGAAACACACACCGGTGTACACCGCTGTACACATTCGTGAAAGAGCGGTGTGTGT 645
QY 523 GTTCGGGCAATGAAGCTGATGATTTGTACCAAGAGTTGATTAATCTGACTGC 582
DB 646 TTCCACAGGATGAAGCTCAACGAGCTGTCAAGATTCAGTTGATTAACAGATACGCG 705
QY 583 TCGGAGAAATTCATTTTGATGAGGCTGAAGAGACACAGATGTGATGTTGGGAGAA 642
DB 706 ACCGCGACCTCTGATTCGATCGAGCATCAACGCGCCACTGATCATCTCTGAGGCGCAAG 765
QY 643 CAATGTGATGATGTGATGATGTAGAGTAGCAAGGCTCTGTGCTCTCAAAAGCT 702
DB 766 AACGTGCTGTCTCGGTACAGGAGATGTGCGCAAGGCTCGCTGAGGCTTTGAGCGC 825
QY 703 CTGGAGCAATTTCTACATTTACCAAAATCGAACCATCTGTGCTGTGAGCGCTGCATG 762
DB 826 CAGGCGCTCTCGCTCAAGAGCTACCAAGCTGACCAATCAACGCTCTTCAAGGCTGTGATG 885
QY 763 GATGGTTACAGGTGTAAAGCTAAATGAATCAATCCGCGAAGTCGATGTGTAATAACT 822

DB 886 GATGGCTACTCTGTGTCTCACCCTGTGATGAGGCGCATTCAGAGAGCGCGCATCTGATACC 945
QY 823 TGCAGAGAAATATAGATGTAGTACACGAGGAGACTTGATGTCAGTGAAGAAACAGTTGT 882
DB 946 GCGACCGGCAACAGGACATCTATTTCTTGAGAGCATGTCAAGATGAAAGATACAGCT 1005
QY 883 ATCGTATGCAATATGAGGCGCATCTCCACACAGAAATGATGTG---ACCAGCTCCGCACT 939
DB 1006 CTGCTGGGCAACATCGGTCTACTTTGATTAAGATGATATGATTCCTCTGTGTGACGCGC 1065
QY 940 CCGGAGCTGACGTGGAGCGAGTACGTTCTCAGGTGACATGTCATCTGCGCAGATGCC 999
DB 1066 GACGACGTCACCCGACACAGATCAACAGGCTCAGCAGAGTTCACTTCTCCACCGGT 1135
QY 1000 AAAGCATTTGCTCCCTCCGCGAGAGGCTGCTACTCATTTGAGCTGCTCCAC---AGT 1096
DB 1126 CGCTCCATCATCTCTGTCTGTCGAGAGGTGCGCTGTGAACCTTGGCAACGCCGAGAC 1185
QY 1057 CCCACCTTTGTTCTGTCCATCAGCAGCAGACAGAGGCTTGGCACTGATGAACCTGTAT 1116
DB 1186 CCATCATTTTGTGATGTCACACTCTTTCGCGCATGACAGCATTTGGCAGATGAACTGTTC 1245
QY 1117 AATGCAACCGAGGCGCATACAGCAGAGATGTACTTGTCTTCTAAGAAATGAGTAA 1176
DB 1246 CA---AAAGCAAGAGCATAGTACAGAGAGTGTACGCTGTGCTAAGTCTGTGACCAA 1302
QY 1177 TACGTTCCAGCTTGCATCTGCCATCTTATATGCTCCACCTTACAGAGCTACAGATGAC 1236
DB 1303 AAGGTGCAACCATTCACAGTTGAGGCTCTCGGCGGTGAGCTACCGAATCAACAGAG 1362
QY 1237 CAAGCAAAATATCTGAGCTCAACAAAATGGCGATTCAAACTAATTTATTAAGATAC 1296
DB 1363 CAGCTGATGATCATCGGCGTTGACGTTGACGAGGCCCATTCAGCGGAGCACTACCGCTAC 1422
RESULT 9
US-09-746-660A-97
; Sequence 97, Application US/0974660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krieger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
; TITLE OF INVENTION: *CORYNEBACTERIUM GLUTAMICUM* GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/746,600A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 97
; LENGTH: 1557
; TYPE: DNA

ORGANISM: Corynebacterium glutamicum
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (101)..(1534)
 OTHER INFORMATION: RXN00132
 US-09-746-660A-97

Query Match 17.4%; Score 225.6; DB 9; Length 1557;
 Best Local Similarity: 56.8%; Pred. No. 4.9e-66;
 Matches 477; Conservative 0; Mismatches 354; Indels 9; Gaps 3;

463 GAGAGAGCGCTGCTGTGCTTACAGCGCTGATACAGCTGCTCCAAAGCTGGAGAACTCTGT 522
 DB 698 GAGAGAAACCAACACCGCTGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 757
 QY 523 GTTCCGCGCATGAGACGTCATGATCTGTTACCAAGAGATTGATTAATCTTACTGC 582
 DB 758 TTCCGACGATGAGACGTCATGACGACGCTGACCAAGTCGCAAGTTGATTAACATACGCG 817
 QY 583 TGCCGAGATTCATTTTGGATGAGCTGAGAGACACAGATGATGTTGGTGGAAA 642
 DB 818 ACCCGGACCTCCCTGATGACGAGCATCAACCGCGCATGATGATGATGATGATGATGATGAT 877
 QY 643 CAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702
 DB 878 AAGCT 937
 QY 703 CTGAGACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 762
 DB 938 CAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
 QY 763 GATGGGTTGAGGTTGATTAAGCTAAATGAGTATCCGCGCAAGTCGATGCTGATTAAT 822
 DB 998 GATGGGTTGAGGTTGATTAAGCTAAATGAGTATCCGCGCAAGTCGATGCTGATTAAT 1057
 QY 823 TGCAGAGAAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 882
 DB 1058 GCGACGCGCAACAGACATCATTTCTGAGCAGATGCTCAAGATGAGAAAGATCAGCT 1117
 QY 883 ATGCTATGCAATATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939
 DB 1118 CTGCTGCGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177
 QY 940 CCGAGCTGACGTTGAGGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 999
 DB 1178 GAGGACGCTGACCGGACACAGATCAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1237
 QY 1000 AAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
 DB 1238 CGCTCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1297
 QY 1057 CCCACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116
 DB 1298 CCATCATTTTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1357
 QY 1117 AATGACCCGAGGCGGATGACAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1176
 DB 1358 CA--AAACCAAGAGACAGTACAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1414
 QY 1177 TAGCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1236
 DB 1415 AAGGTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1474
 QY 1237 CAGGCAAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1296
 DB 1475 CAGGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1534

RESULT 10
 US-10-021-121-1/c
 Sequence 1, Application US/10021121
 Patent No. US2002014244A1
 GENERAL INFORMATION:

APPLICANT: Caras, Ingrid W
 TITLE OF INVENTION: A2-1 Neurotrophic Factor
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 1 DNA Way
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Winpatlin (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/021,121
 FILING DATE: 06-Dec-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/635,130
 FILING DATE: 19-Mar-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Torchia, PhD, Timothy E.
 REGISTRATION NUMBER: 36,700
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650/225-8674
 TELEFAX: 650/952-9881
 INFORMATION FOR SEQ. ID NO. 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1877 base pairs
 TYPE: Nucleic Acid
 STRANDEDNESS: Double
 TOPOLOGY: Linear
 FEATURE:
 NAME/KEY: Extra Cellular Domain
 LOCATION: 244-899
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 FEATURE:
 NAME/KEY: Transmembrane Domain
 LOCATION: 901-978
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 FEATURE:
 NAME/KEY: signal peptide
 LOCATION: 244-321
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-10-021-121-1
 Query Match 14.7%; Score 190; DB 12; Length 1877;
 Best Local Similarity 100.0%; Pred. No. 7.9e-54;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 735 CCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 794
 DB 1446 CCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1387
 QY 795 CATCCGCGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 854
 DB 1386 CATCCGCGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1327
 QY 855 GCACTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 914
 DB 1326 GCACTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267
 QY 915 AATGATGCTG 924
 DB 1266 AATGATGCTG 1257

RESULT 11

US-09-746-660A-101
; Sequence 101, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberer, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
; TITLE OF INVENTION: CORNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/746,660A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO.101
; LENGTH: 1396
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1396)
; OTHER INFORMATION: FRXA01371
US-09-746-660A-101

Query Match 14.4%; Score 186; DB 9; Length 1396;
Best Local Similarity 57.3%; Pred. No. 1.5e-52;
Matches 377; Conservative 0; Mismatches 275; Indels 6; Gaps 2;

QY 463 GAAGAGAGCTGACTGCTGTCACAGGCTGTATCAAGCTCTCCAAAGCTGGAGAGCTCTGT 522
DB 698 GAGGAAACACACCGGCTGCACCGCTGTACCACTTCCTGAAGAGCGCTGCTGCT 757
QY 523 GTCCGGCAGTGAAGCTGATTCGTTACCAAGAGTTGATTAAGTCTGACTGC 582
DB 758 TTCCAGAGATGAAAGCTGACAGAGCTGTACCAAGTTGATTAAGAGAGCGC 817
QY 583 TGCCGAGATTCATTTGATGAGCTGTGAAGAGACCAAGATGTGATTTGGTGGAAA 642
DB 818 ACCGGCAGCTCCGATGACAGGAGCATCAACGGCGCACTGATGATGGGGGGCAAG 877
QY 643 CAATGGTGTGTGTGCTATGTGAGTAGGCAAGGGCTCTGTCTCTCAAACT 702
DB 878 AACGTGTGTGTGTGCTATGTGAGTAGGCAAGGGCTCTGTCTCTCAAACT 937
QY 703 CTTCGAGATTCATTTGATGAGCTGTGAAGAGACCAAGCTGTGCTGAGGCGCGATG 762
DB 938 CAGGGCGCTCCGCTCAAGAGTACAGAGCTGACCAATCAACGCTCTCAAGGCTGTATG 997
QY 763 GATGGCTTCAGGCTGTAAAGCTTAATGAACTATCCGCAAGTGTGATGTATATACT 822
DB 998 GATGGCTTCAGGCTGTAAAGCTGTATGATGAGCAATCGAAGAGCGCGAGATGTATATACC 1057

RESULT 12

US-09-738-626-838/c
; Sequence 838, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 838
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-838

Query Match 13.4%; Score 174; DB 9; Length 720;
Best Local Similarity 61.0%; Pred. No. 1.3e-48;
Matches 282; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 463 GAAGAGAGCTGACTGCTGTCACAGGCTGTATCAAGCTCTCCAAAGCTGGAGAGCTCTGT 522
DB 465 GAGGAAACACACCGGCTGCACCGCTGTACCACTTCCTGAAGAGCGCTGCTGCT 406
QY 523 GTCCGGCAGTGAAGCTGATTCGTTACCAAGAGTTGATTAAGTCTGACTGC 582
DB 405 TTCCAGAGATGAAAGCTGACAGAGCTGTACCAAGTTGATTAAGAGAGCGC 346
QY 583 TGCCGAGATTCATTTGATGAGCTGTGAAGAGACCAAGATGTGATTTGGTGGAAA 642
DB 345 ACCGGCAGCTCCGCTGATGAGGAGCATCAACGGCGCACTGATGATGGGGGGCAAG 286
QY 643 CAATGGTGTGTGTGCTATGTGAGTAGGCAAGGGCTCTGTCTCTCAAACT 702
DB 285 AACGTGTGTGTGTGCTATGTGAGTAGGCAAGGGCTCTGTCTCTCAAACT 226


```

Db 178798 GACCGAGATGATCCCATCTGTCCCTTCAGGCTCTCATGTGAAGGCGCTTACAGTTCTGAC 178739
QY 783 GCTAAATGAGTCAATCCGCGCACTGATGCTGAATTAATCTGCACAGAGAAATAGAAATGT 842
Db 178738 CTTGGAGGATGTGTCTTGAGAGCTGATATCTTGTCCACCAACCGGTAACAGAGCAT 178679
QY 843 AGTGACACGGGAGACATTTGGATGCGATGAAAAACAGTTGATGTATGCAATTTGGGCCA 902
Db 178678 CACCAATGTTGACCATATGAGAAATGAAAGAACATGCCATTTGTTCAACATTTGCTCA 178619
QY 903 CTCACACACAGAAATGATGATGACACAGCTCCGCACT---CCGAGCTGACGTGGAGCG 959
Db 178618 CTTTGACATATGATGACATGCTGTGGCGTGGAGAACTACCCCGGCTGAAGGCAATCAC 178559
QY 960 AGTAGCTTCTAGGTGAGACCATGCTCATCTGGCCAGATGGCAACAGATTGTC---CTCCT 1016
Db 178558 CATCAAGCCCAACATGACAGATGGGTCTTCCCTGAGACCAACACCGGTATCATTTGCTT 178499
QY 1017 GGCAGAGGGTCTGTACTCAATTTGAGCTGCTCCACAGTCCAC---TTTGTCTGTC 1073
Db 178498 GGGTGAAGGGTCGATGATGAACTGGGATGCGCACTGGACACCCAGTTTGTGATGTC 178439
QY 1074 CATCAAGCCCAACACAGGCTTGGCAGTATAGAACTATATATACACCCG---AGGG 1130
Db 178438 CTGCTCTTTCACCAACAGGCTCATCTGCTCAGCTTGAGTTTGGAGAGAGAGATACCG 178379
QY 1131 GCGATACAGAGAGATGTACTGCTTCTCTAAGAAAAATGATGATACGTTGGCAGCTT 1190
Db 178378 CAGTAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 178319
QY 1191 GCATCTGCTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1250
Db 178318 TCACCTGGGCAACCTGAGCTTAAGCTGACCCAGCTTAGCAAGTCCAGAGCTGATTAAT 178259
QY 1251 GGGACTCAACAAAATGGGCGCATTAACAACTAATTAATTAACAGATAC 1296
Db 178258 CAGTGTGCTGTTGAGGGGTGCATACAAAGCTGCTCACTACAGGTAC 178213

RESULT 15
US-09-960-352-4421
; Sequence 4421, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengding
; APPLICANT: Byatt, John C.
; APPLICANT: Mathias, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960/352
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4421
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 19-LIB34-025-Q1-E1-E3
US-09-960-352-4421

Query Match 13.08; Score 168.2; DB 10; Length 433.
Best Local Similarity 64.08; P-Val 8.9e-47;
Matches 254; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 193 GAGACACTGTGTGCGGCGGCTAGTGGCGGCTGCTGTTGTAACATCTACTCACT 252
Db 37 GAGACCTCTGTTGCGGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGT 96
QY 253 CAGATAGATAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312
Db 97 CAGAGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 156

```

```

QY 313 GAGTCAAGATGATCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 372
Db 157 GAACGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 216
QY 373 GCGACATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 432
Db 217 CTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276
QY 433 AACGTGTTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
Db 277 CAGCTCTGTCAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 336
QY 493 TATCAGTCTCCAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 552
Db 337 TACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 396
QY 553 ACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 589
Db 397 ACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433

```

Search completed: April 21, 2003, 07:31:10
 Job time : 854.419 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 23:31:25 ; Search time 1749.62 Seconds
(without alignments)
11996.532 Million cell updates/sec

Title: US-09-782-051-1_COPY_549_1844

Perfect score: 1296
Sequence: 1 tccaaaggcagcagcaattt.....aacctaattacagataac 1296

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	845.6	65.2	1042	13	BM449470 AGENCOURT
2	834.8	64.4	3234	11	AK014539 Mus muscu
3	775.6	59.8	871	12	BG708016 602671341
4	732.2	56.5	894	14	BQ896777 AGENCOURT
5	730.4	56.4	748	9	AU124992 AU124992
6	720.4	55.6	933	14	BQ889173 AGENCOURT

7	716.4	55.3	747	12	BG697210 602660444
8	716.4	55.3	832	10	BE270773 BE270773
9	698.2	53.9	751	9	AU133711 AU133711
10	692.4	53.4	745	9	AU131962 AU131962
11	687.8	53.1	691	14	BM837838 K-EST0113
12	673	51.9	1098	13	BM454352 AGENCOURT
13	672.4	51.9	956	12	BF792750 602253481
14	666.2	51.4	938	9	AL527928 AL527928
15	661.8	51.1	930	12	BG751929 602371492
16	655.8	50.6	764	9	AU126038 AU126038
17	648.4	50.0	764	12	BG501360 602547990
18	644	49.7	1379	14	U69200 U69200
19	635.8	49.1	773	14	BQ745712 UT-M-EMO-
20	629.8	48.6	927	10	BE560874 601346168
21	627.4	48.4	629	14	BM786844 K-EST0065
22	622.6	48.0	827	12	BE777877 601463641
23	621	47.9	839	14	BQ212518 AGENCOURT
24	613	47.3	614	14	BM837761 K-EST0113
25	603.8	46.6	819	10	BE542285 601067027
26	599.8	46.3	969	12	BG259448 602378544
27	597.6	46.1	639	10	AM958506 EST370576
28	593.8	45.8	643	14	BM955818 UT-H-EDO-
29	592	45.7	598	14	BQ614546 1X98402.Y
30	588.8	45.4	664	10	AV751644 AV751644
31	588.8	45.4	947	12	BF792477 602253566
32	587.6	45.3	1124	14	BQ062203 AGENCOURT
33	586.4	45.2	867	13	BM453920 AGENCOURT
34	585	45.1	1021	14	BQ059373 AGENCOURT
35	582.2	44.9	1002	12	BE900605 601673502
36	578	44.6	590	10	BE514092 601315914
37	575.6	44.4	665	12	BG898694 HOA4-1-A-
38	572.6	44.2	672	12	BF342864 602015152
39	565	43.6	844	10	BE536849 601064852
40	564.4	43.5	731	12	BG826655 602748977
41	562	43.4	700	9	AU133884 AU133884
42	561.4	43.3	975	10	BE501484 601346553
43	558.8	43.1	603	12	BG244104 602358170
44	551.6	42.6	616	12	BF783516 602111292
45	542.2	41.8	858	9	AL554850 AL554850

ALIGNMENTS

RESULT 1
LOCUS BM449470
DEFINITION AGENCOURT_6400861 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5493677
ACCESSION BM449470
VERSION BM449470.1 GI:18498510
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/ (Bases 1 to 1042)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LLM2117 row: m column: 06
High quality sequence stop: 694.
Location/Qualifiers 1..1042

FEATURES
source

QY	1198	CCATATTTGAGTGGCCACCTTACAGAGCTGCAGATGACCAAAATATGTGGAGCT	1257
Db	1787	CCGACCTTTGATGCCCACTGACAGAGCTTACAGATGACCAAGCAAGTATCTGGGACTC	1846
QY	1238	AACAAAATGGGCGCATTCAAACCTATATATATACAGATA	1295
Db	1847	AACAAGATGAGACCTTCAAGCCTTAAGTACATACAGTA	1884
RESULT 3			
LOCUS	Bg708016		
DEFINITION	Bg708016	871 bp	linear
ACCESSION	Bg708016		EST 07-MAY-2001
VERSION	Bg708016.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: c96pbp@remail.nih.gov		
	Tissue Procurement: Miklos Palcovits, M.D., Ph.D.		
	cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki		
	Toshiyuki and Piero Carninci (RIKEN)		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.llnl.gov		
	Plate: LLM10674	row: b	column: 11
	High quality sequence stop: 839.		
FEATURES			
source	location/Qualifiers		
	1..871		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_image="4793770"		
	/clone_id="NIH_MGC_96"		
	/tissue_type="hypothalamus"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pBluescript (modified		
	pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag		
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTVN-3',		
	size-selected for average insert size 2.3 kb and		
	normalized to R0F 5. This is a primary library enriched		
	for full-length clones and constructed using the		
	Cap-trapper method (Carninci, in preparation). Library		
	constructed by M. Brownstein (NIH/NHGRI, National		
	Institutes of Health). Note: this is a NIH-MGC Library."		
BASE COUNT	238 a	190 c	235 g
ORIGIN			207 t
			1 others
Query Match	59.8%	Score 775.6;	DB 12; Length 871;
Best Local Similarity	98.5%	Prod. NO. 1.9e-28;	
Matches 803; Conservative	0;	Mismatches 10;	Indels 2; Gaps 2;
QY	360	GGATGGGTGGCGAGCGCAACATGATCCTGGATGATGGGAGAGACTTAACCACTGGGTTTA	419
Db	5	GGATGGGTGGCGAGCGCAACATGATCCTGGATGATGGGAGAGACTTAACCACTGGGTTTA	64
QY	420	TGAGAGTATCCAAAGCTGTTTAAGAGATCCGAGCATTTGGAAGAGAGCGTACTGG	479
Db	63	TGAGAGTATCCAAAGCTGTTTAAGAGATCCGAGCATTTGGAAGAGAGCGTACTGG	124
QY	480	TTTTCACAGGCTGTATCAGCTCTCCAAAGCGGGAGAGCTCTGTCTCGGCCCATGAAGT	539
Db	125	TTTTCACAGGCTGTATCAGCTCTCCAAAGCGGGAGAGCTCTGTCTCGGCCCATGAAGT	184

QY	540	CAATATTTCTGTACCAACAGAAATTTGATTAATCTGTAGCTGCTCCGGAATTCATTTT	599
Db	185	CATATATTTCTGTACCAACAGAAATTTGATTAATCTGTAGCTGCTCCGGAATTCATTTT	244
QY	600	GGATGGCCTGAAAGAGACCAAGATGTGATGTTTGGTGGGAAACAAGTGGTGTGG	659
Db	245	GGATGGCCTGAAAGAGACCAAGATGTGATGTTTGGTGGGAAACAAGTGGTGTGG	304
QY	660	CTATGTGAGGTAGCCAAAGGCTGCTGTGTCTCTCAAGCTCTTTGAGCATTTGTCTA	719
Db	305	CTATGTGAGGTAGCCAAAGGCTGCTGTGTCTCTCAAGCTCTTTGAGCATTTGTCTA	364
QY	720	CATTACCGAAATCGAACCCTATCTGTCTGTGAGGCTTCAGATGAAATGAGGTGTGT	779
Db	365	CATTACCGAAATCGAACCCTATCTGTCTGTGAGGCTTCAGATGAAATGAGGTGTGT	424
QY	780	AAAGCTAAATGAAATGATCCGCAAGTGTGATGCTGTAATTAATTCACAGAAATAGAA	839
Db	425	AAAGCTAAATGAAATGATCCGCAAGTGTGATGCTGTAATTAATTCACAGAAATAGAA	484
QY	840	TCTAGTGACACGGGAGCACTTGGATTCGATGAAAAACAGTTGTATTCGTAATATGG	899
Db	485	TCTAGTGACACGGGAGCACTTGGATTCGATGAAAAACAGTTGTATTCGTAATATGG	544
QY	900	CCAGTCGAACAGAGAAATGATGTGACCAAGCTCCGCACTCCGGAAGCTACGTGGAGCG	959
Db	545	CCAGTCGAACAGAGAAATGATGTGACCAAGCTCCGCACTCCGGAAGCTACGTGGAGCG	604
QY	960	AGTACGTTCTCAGGTGAGACCATGTGATGCGCAAGTGGCAAAAGATTTGCTCTGGC	1019
Db	605	AGTACGTTCTCAGGTGAGACCATGTGATGCGCAAGTGGCAAAAGATTTGCTCTGGC	664
QY	1020	AAGAGGTGCTACACATTAATGAGTGTGCGCACAGTCCACCTTGTCTCTCATGAC	1079
Db	665	AAGAGGTGCTACACATTAATGAGTGTGCGCACAGTCCACCTTGTCTCTCATGAC	724
QY	1080	AGCCACACACAGGCTTTGGCACT-GATGAAGCTCTAATATGACACCGAGGGGAGTACA	1138
Db	725	AGCCACACACAGGCTTTGGCACTGATGAAGCTCTAATATGACACCGAGGGGAGTACA	783
QY	1138	AGCAGAGTGTACTTCTCTTAAGAAATGATGAT 1173	
Db	784	AGCAGAGTGTACTTCTCTTAAGAAATGATGAT 818	
RESULT 4			
BO896777			
LOCUS			
DEFINITION			
AGENCOURT_8715624 NIH.MGC.101 Homo sapiens CDNA clone IMAGE:645821			
5', mRNA sequence.			
ACCESSION			
BO896777			
VERSION			
BO896777.1 GI:22288791			
KEYWORDS			
EST.			
SOURCE			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
1 (bases 1 to 894)			
NIH-MGC http://mhc.ncl.nih.gov/			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
Contact: Robert Strusberg, Ph.D.			
Email: cgrabs-remail.nih.gov			
Tissue Procurement: ATCC			
CDNA Library Preparation: Rudin Laboratory			
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Agencourt Bioscience Corporation			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: L12M2608 row: d column: 14			
High quality sequence stop:700.			


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6185747"
/clone_11b="lupsk1_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/notes="Vector: pCMV-Sport6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
directionally cloned using the following adaptors:
5'-TGACGCCACGGCGTCGG-3' and
5'-GACTGATGTTTGATGTCGACGAGGCGCCGCTT(15)-3'. Size selected >

```

OY	113	GTGCTCAGGGGAGAGAGCCCTGGCTGGGTGCTAAATAATAGGCGCTACACATCAAG	172
Db	1	GTGCTCAGGGGAGAGAGCCCTGGCTGGGTGCTAAATAATAGGCGCTACACATCAAG	60
OY	173	CCGACAGACGGGTGTTGATTGAGACACTGTGTCCCTGGGGGCTCAGTCCGCTGGTCTG	232
Db	61	CCGACAGACGGGTGTTGATTGAGACACTGTGTCCCTGGGGGCTCAGTCCGCTGGTCTG	120
OY	233	CTTGTACATCTACACACTCGAATGAAATGCTGCACACTGGCTGAGGCTGGAGTTG	292
Db	121	CTTGTACATCTACACACTCGAATGAAATGCTGCACACTGGCTGAGGCTGGAGTTG	180
OY	293	CAGTGTGCTTGGAAAGGGCGAGTCAGAAAGATGCTTCTGTGTGTATGACCCGTGTG	352
Db	181	CAGTGTGCTTGGAAAGGGCGAGTCAGAAAGATGCTTCTGTGTGTATGACCCGTGTG	240
OY	353	TGAACATGATGATGGGTGGCGAGCCCAACATGATCCTGGATGATGGGGACATTACCCT	412
Db	241	TGAACATGATGATGGGTGGCGAGCCCAACATGATCCTGGATGATGGGGAGCTTACCCT	300
OY	413	GGGTTTATAGAAATATCCAAACGCTGTTTAAAGATCCGAGGCAATTGGAAAGAGCG	472
Db	301	GGGTTTATAGAAATATCCAAACGCTGTTTAAAGATCCGAGGCAATTGGAAAGAGCG	360
OY	473	TGACTGCTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGGCTGTGTTCCGGCCA	532
Db	361	TGACTGCTTCACAGGCTGTATCAGCTCTCCAAAGCTGGAGGCTGTGTTCCGGCCA	420
OY	533	TGAACGTAAATGATCTCTGTACCAACAGAGTTGATATCTTGTACTGCTCCGAGAT	592
Db	421	TGAACGTAAATGATCTCTGTACCAACAGAGTTGATATCTTGTACTGCTCCGAGAT	480
OY	593	CCATTGTGATGGGCTGTAAGAGGACACAGATGATGTTGGTGGAAACAAGTGCTGG	652
Db	481	CCATTGTGATGGGCTGTAAGAGGACACAGATGATGTTGGTGGAAACAAGTGCTGG	540
OY	653	TGTGTGCTATGCTAGAGTAGAGCAAGGCGCTGTGTGCTCTCCAAAGCTTTGGAGCA	712
Db	541	TGTGTGCTATGCTAGAGTAGAGCAAGGCGCTGTGTGCTCTCCAAAGCTTTGGAGCA	600
OY	713	TTGTCTACATTACCGAAATCGAACCCCATCTGTGCTCTCAGGCGCTGCATGATGGG--TT	770
Db	601	TTGTCTACATTACCGAAATCGAACCCCATCTGTGCTCTCAGGCGCTGCATGATGGG	660
OY	771	CAGGTGTGTTAAAGCTAAATGAATGATCATCCGCGAGTGCATGTCTGAATTACTTCACAG	830
Db	661	AGGGTGGGTAAAGCTAAATGAATGATCATCCGCGAGTGCATGTCTGAATTACTTCACAG	720
OY	831	AAATAGAAATAGTGAACAC--GGGAGACACTGGATGCGATG--AAAAACAGTTGATC	885
Db	721	AAATAGAAATAGTGAACACCGGAGACACTTGGATGCGATGCGAAAAACAGTTGATC	778

RESULT 7	B6697210	747 bp	mRNA	linear	EST 07-MAY-2001
LOCUS	B6697210				
DEFINITION	602666444F1 NC1_CGAP_Skn3 Homo sapiens cDNA IMAGE:4803515 5', mRNA sequence.				
ACCESSION	B6697210				
VERSION	B6697210.1	GI:13963166			
KEYWORDS	EST.				

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 747)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1069 row: h column: 12
High quality sequence stop: 747.
Location/Qualifiers
1. 747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4803515"
/lab_host="NCI CGAP_Skn3"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT 180 a 152 c 232 g 183 t
ORIGIN

Query Match 55.3%; Score 716.4; DB 12; Length 747;
Best Local Similarity 99.6%; Pred. No. 2.6e-208;
Matches 739; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 34 ATCAAGCAGGAGATTTGGACGCCGGAGATTGACAGACCAACATCTGCTGCT 93
DB 1 ATCAAGCAGGAGAA-TTGGACGCCGGAGATTGAGATTGACAGACCAACATCTGCTGCT 59
QY 94 CTGATTTACTCAGGAAACGCTCAGGAGGAGAGACCCCTGGCTGGCTAAATAGG 153
DB 60 CTGATTTACTCAGGAAACGCTCAGGAGGAGAGACCCCTGGCTGGCTAAATAGG 119
QY 154 GCGGTACACACATCACAGCCAGACGCGGTGATTGAGACACTCTGTCCTGGGG 213
DB 120 GCGGTACACACATCACAGCCAGACGCGGTGATTGAGACACTCTGTCCTGGGG 179
QY 214 GCTAGTGCCTGCTGCTGCTGTTTAACATCTCACTCACTCAATGAATGAGTCTGACGA 273
DB 180 GCTAGTGCCTGCTGCTGCTGTTTAACATCTCACTCACTCAATGAATGAGTCTGACGA 239
QY 274 CTGCTGAGGCTGAGGTTGAGGTTGCTGTTGAAAGGCGAGTCAAGATGACTTCGG 333
DB 240 CTGCTGAGGCTGAGGTTGAGGTTGCTGTTGAAAGGCGAGTCAAGATGACTTCGG 299
QY 334 TGGTGTATTGACCCGCTGTGTGAACATGATGGGTGGAGGCCAATGATCCTGGATGAT 393
DB 300 TGGTGTATTGACCCGCTGTGTGAACATGATGGGTGGAGGCCAATGATCCTGGATGAT 359
QY 394 GGGGAGAGACTTAACCCACTGGGTTTAAAGATATCCAAA-CGTGTTAAGAGATCCG 452
DB 360 GGGGAGAGACTTAACCCACTGGGTTTAAAGATATCCAAAAGCGTTAAGAGATCCG 419
QY 453 AGGCAATGTGAGAGAGCGTGAAGTGTGACAGGCTGATGAGCTCTCCAAAGCTGG 512
DB 420 AGGCAATGTGAGAGAGCGTGAAGTGTGACAGGCTGATGAGCTCTCCAAAGCTGG 479
QY 513 GAACCTGTGCTCCGGCCATGAACGTCATGATTTCTTACCAAGAGAGATTGATTA 572
DB 480 GAACCTGTGCTCCGGCCATGAACGTCATGATTTCTTACCAAGAGAGATTGATTA 539

QY 573 CTGTACTGCTGCCGAGATTCATTTTGATGGCCCTGAAGAGACCAAGATGTATGTT 632
DB 540 CTGTACTGCTGCCGAGATTCATTTTGATGGCCCTGAAGAGACCAAGATGTATGTT 599
QY 633 TGGTGGAAACAAGTGTGTGTGCTATGATGTAGTAGGCAAGGCTGCTGCTCG 692
DB 600 TGGTGGAAACAAGTGTGTGTGCTATGATGTAGTAGGCAAGGCTGCTGCTCG 659
QY 693 TCTCAAGCTCTTGGAGCAATGTCTTACATTAACCAAGATGACCCCATCTGCTGTGA 752
DB 660 TCTCAAGCTCTTGGAGCAATGTCTTACATTAACCAAGATGACCCCATCTGCTGTGA 719
QY 753 GGCCTGATGATGATGGCTTACG 774
DB 720 GGCCTGATGATGATGGCTTACG 741

RESULT 8
BE270773 832 bp mRNA linear EST 13-JUL-2000
LOCUS 600943931P1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:296598 5',
DEFINITION mRNA sequence.
ACCESSION BE270773
VERSION BE270773.1 GI:9144416
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 832)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM69 row: n column: 07
High quality sequence start: 2
High quality sequence stop: 767.
Location/Qualifiers
1. 832
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:296598"
/clone_id="NIH_MGC_8"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 236 a 192 c 202 g 201 t
ORIGIN

Query Match 55.3%; Score 716.4; DB 10; Length 832;
Best Local Similarity 99.5%; Pred. No. 2.8e-208;
Matches 750; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

QY 545 ATTCTGTACCAACAAGAGATTGATTAATCTGTGCTGCCGAGATTCATTTGATG 604
DB 13 ATTCTGTACCAACAAGAGATTGATTAATCTGTGCTGCCGAGATTCATTTGATG 71
QY 605 GCGTAAAGAGACCAAGATGTGATGTTGGTGGAAACAGAGTGTGTGCTGCTATG 664

```

Db 72 GCTGTAAGAGGACACAGATGTGATGTTGGGAAACAAGTGGTGTGTGTGCTATG 131
QY 665 GTGAGGTAGGCAAGGGCTGTGTGTCTCTCAAAAGCTCTGGAGCAATGTCTACATA 724
Db 132 GTGAGGTAGGCAAGGGCTGTGTGTCTCTCAAAAGCTCTGGAGCAATGTCTACATA 191
QY 725 CCGAAATGACCCCATCTGTGTCTGTGACAGGCTGCATGATGGGTTCAGGGGTAAAGC 784
Db 192 CCGAAATGACCCCATCTGTGTCTGTGACAGGCTGCATGATGGGTTCAGGGGTAAAGC 251
QY 785 TAAATGAATCATCCGGAAGTGCATGTCTATTAATACCTGCACAGAAATAGAATGAG 844
Db 252 TAAATGAATCATCCGGAAGTGCATGTCTATTAATACCTGCACAGAAATAGAATGAG 311
QY 845 TGACACGGGAGACCTTGTGATGCAATGAAAAACAGTTGTATGCAATATAGGCCACT 904
Db 312 TGACACGGGAGACCTTGTGATGCAATGAAAAACAGTTGTATGCAATATAGGCCACT 371
QY 905 CCAACACAGAAATGATGTGACACAGCTCCGACCTCCGAGCTGCAGTGGAGAGAGTAC 964
Db 372 CCAACACAGAAATGATGTGACACAGCTCCGACCTCCGAGCTGCAGTGGAGAGAGTAC 431
QY 965 GTTCTCAGGTGAGACCATGTCTATGTGCGCAGATGGCAAGAGTTGTCTCTGCGAGAG 1024
Db 432 GTTCTCAGGTGAGACCATGTCTATGTGCGCAGATGGCAAGAGTTGTCTCTGCGAGAG 491
QY 1035 GTTCTCAGGTGAGACCATGTGAGCTGTCCACAGTCCCACTTTTGTGTGCAATACAGCA 1084
Db 492 GTTCTCAGGTGAGACCATGTGAGCTGTCCACAGTCCCACTTTTGTGTGCAATACAGCA 551
QY 1085 CAACACAGGCTTGGCAGCTGATAGAACTTATATGACACCCGAGGCGCATACAGCAG 1144
Db 552 CAACACAGGCTTGGCAGCTGATAGAACTTATATGACACCCGAGGCGCATACAGCAG 611
QY 1145 ATGTCTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGAT 1204
Db 612 ATGTCTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGAT 671
QY 1205 TTGATGCCACCTTACAGAGCTGACAGATGACCAAGC-AAATATCTGGGATCTCAACAA 1263
Db 672 TTGATGCCACCTTACAGAGCTGACAGATGACCAAGC-AAATATCTGGGATCTCAACAA 731
QY 1264 AATGGGCAATC-AAACCTAATTAATTAACAGATAC 1296
Db 732 AATGGGCAATC-AAACCTAATTAATTAACAGATAC 765

RESULT 9
LOCUS AU133711 751 bp mRNA linear EST 01-AUG-2002
DEFINITION AU133711 OVARC1 Homo sapiens cDNA clone OVARC100494 5', mRNA
sequence.
ACCESSION AU133711
VERSION AU133711.1 GI:10994250
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 751)
Ota.T., Nishikawa.T., Suzuki.Y., Ishii.S., Saito.K., Kawai.Y.,
Yamamoto.C.T., Wakamatsu.A., Nakamura.Y., Nagai.T., Sugano.S. and
Iboga.T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isegai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix

```

```

FEATURES
Source
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="OVARC100494"
  /clone_lib="OVARC1"
  /tissue_type="ovary, tumor tissue"
  /note="Vector: pMR18FL3"
BASE COUNT 192 a 149 c 222 g 184 t 4 others
ORIGIN
Query Match 53.9% Score 698.2; DB 9; Length 751;
Best Local Similarity 99.0%; Pred. No. 1e-202;
Matches 711; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 101 CACTCAGGAAACGTCTCAGGGGAGAAACCTTGGCTGTGCTAAATATAGTGGCTGTA 160
Db 1 CACTCAGGAAACGTCTCAGGGGAGAAACCTTGGCTGTGCTAAATATAGTGGCTGTA 60
QY 161 CACACATCACAGCCAGACAGCGGTGTATGAGACACTCTGTGCCCCGGGGCTCACT 220
Db 61 CACACATCACAGCCAGACAGCGGTGTATGAGACACTCTGTGCCCCGGGGCTCACT 120
QY 221 GCCGCGGTCTGTGTATCATCTACTCACTCACTAGAAATAGATGAGAGCACTGGCTG 280
Db 121 GCCGCGGTCTGTGTATCATCTACTCACTCACTAGAAATAGATGAGAGCACTGGCTG 180
QY 281 AGGCTGAGATTCAGATGTTCCGTTGGAAAGGCGAGTCAGAAATGACTTGTGTGTGTA 340
Db 181 AGGCTGAGATTCAGATGTTCCGTTGGAAAGGCGAGTCAGAAATGACTTGTGTGTGTA 240
QY 341 TTGACCGCTGTGTGAACATGATGGTGGCAGAGCCCAACATGATCCTCGATGATGGGGAG 400
Db 241 TTGACCGCTGTGTGAACATGATGGTGGCAGAGCCCAACATGATCCTCGATGATGGGGAG 300
QY 401 ACTTAACCCAGTGGGTTTAAAGATATCCAAAGCTGTTAAGAAATGCCAGGATG 460
Db 301 ACTTAACCCAGTGGGTTTAAAGATATCCAAAGCTGTTAAGAAATGCCAGGATG 360
QY 461 TGAAGAGAGCGTGAAGTGTGTTACAGAGCTGTATGACGCTCTCCAAAGCTGGAGACTG 520
Db 361 TGAAGAGAGCGTGAAGTGTGTTACAGAGCTGTATGACGCTCTCCAAAGCTGGAGACTG 420
QY 521 GTGTCCGGCCATGAAGCTGATGATGTTGTTACAAAGAAAGTTGATTAATCTTACT 580
Db 421 GTGTCCGGCCATGAAGCTGATGATGTTGTTACAAAGAAAGTTGATTAATCTTACT 480
QY 581 GCTGCCGAATATCCATTTGGATGGGCTGAAGAGGACACAGATGATGTTGGTGGGA 640
Db 481 GCTGCCGAATATCCATTTGGATGGGCTGAAGAGGACACAGATGATGTTGGTGGGA 540
QY 641 AACAAAGTGTGTGTGTGCTATGATGAGGTAGCAAGAGGCTCTGTGCTCTCAAG 700
Db 541 AACAAAGTGTGTGTGTGCTATGATGAGGTAGCAAGAGGCTCTGTGCTCTCAAG 600
QY 701 CTTGTGAGCAATGTCTACATTAACGAAATGCAACCCCATCTGTCTGACAGCCCTGCA 760
Db 601 CTTGTGAGCAATGTCTACATTAACGAAATGCAACCCCATCTGTCTGACAGCCCTGCA 660
QY 761 TGGATGGGTTCAGGGT-GGTAAGCTAAATGAAGTCAATCCGGAAGTCATGTCTGTA 817
Db 661 TGGATGGGTTCAGGGTGGTGGTAAGCTAAATGAAGTCAATCCGGAAGTCATGTCTGTA 718

RESULT 10
LOCUS AU131962 745 bp mRNA linear EST 01-AUG-2002
DEFINITION AU131962 NT2RP3 Homo sapiens cDNA clone NT2RP3003545 5', mRNA
sequence.
ACCESSION AU131962

```

VERSION	EST.	Accession
AU31962.1	GI:10992316	
KEYWORDS		
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eunariya, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 745)	
AUTHORS	Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Salto, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.	
TITLE	HRI human cDNA project	
JOURNAL	Unpublished (2000)	
COMMENT	Contact: Takeo Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute. Location/Qualifiers 1..745 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="NT2RP303545" /clone_11b="NT2RP3" /cell_type="teratocarcinoma" /cell_line="NT2" /note="Vector: pME18SF12; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction" BASE COUNT 193 a 149 c 214 g 186 t 3 others ORIGIN	
FEATURES		
source		
Query Match	53.4%; Score 692.4; DB 9; Length 745;	
Best Local Similarity	99.6%; Pred. No. 6e-201;	
Matches 704; Conservative	0; Mismatches 2; Indels 1; Gaps 1;	
Db	1 TCTGCTGTATACATCTACTCACTCAGATGAGTAGTGTGACGACACTGGCTGGA 288	
Qy	229 TCTGCTGTATACATCTACTCACTCAGATGAGTAGTGTGACGACACTGGCTGGA 288	
Db	1 TCTGCTGTATACATCTACTCACTCAGATGAGTAGTGTGACGACACTGGCTGGA 60	
Qy	289 GTTGACAGCTTCGCTTGGAAGGGCGAGTCAGAAATGACTTCTGGTGTGATGACCC 348	
Db	61 GTTGACAGCTTCGCTTGGAAGGGCGAGTCAGAAATGACTTCTGGTGTGATGACCC 120	
Qy	349 TGTGGAACATGAGTGGTGGGAGGCCAACAATGATCCGTGATGATGGGGGAGACTTAAC 408	
Db	121 TGTGGAACATGAGTGGTGGGAGGCCAACAATGATCCGTGATGATGGGGGAGACTTAAC 180	
Qy	409 CACTGGGTTTATAGAAATATCAAAACCTGTTTAAAGAAATCCGAGCATTTGTGAAG 468	
Db	181 CACTGGGTTTATAGAAATATCAAAACCTGTTTAAAGAAATCCGAGCATTTGTGAAG 240	
Qy	469 AGCGTACTGTTGTCACAGCGCTGATCAGCTCTCCAAAGCTGGGAGCTGTGTTCCG 528	
Db	241 AGCGTACTGTTGTCACAGCGCTGATCAGCTCTCCAAAGCTGGGAGCTGTGTTCCG 300	
Qy	529 GCCATGAACGTCATGATCTCTGTTACCAAAAGAAAGTTGTTAACTTTACTGCTGCCGA 588	
Db	301 GCCATGAACGTCATGATCTCTGTTACCAAAAGAAAGTTGTTAACTTTACTGCTGCCGA 360	
Qy	589 GAATCATTTTGGATGGGCTGAAGAGGACCAACAGTGTGATGTTGGTGGGAACAATG 648	
Db	361 GAATCATTTTGGATGGGCTGAAGAGGACCAACAGTGTGATGTTGGTGGGAACAATG 420	
Qy	649 GTGCTGTGCTGATGATGAGGTAGGCAAGAGGCTGCTGTGCTCTCAAGACTTTTGA 708	
Db	421 GTGCTGTGCTGATGATGAGGTAGGCAAGAGGCTGCTGTGCTCTCAAGACTTTTGA 480	
Qy	709 GCAATGCTCTACATTTACCGAAATGACCCCATCTGTGCTCTGCAGGCTGCATGAGTGG 768	

Db	481	GCAGATGTCTAATTCGAGAAATCGACCCCATCTGTGCTCGACGGCCCTGGATGATGGG	540
Qy	769	TTCCAGGCTGTGAACCTAAATGATGCATCCCGCAAGTCGATGCTGTAATACTGCACCA	828
Db	541	TTCCAGGCTGTGAACCTAAATGATGCATCCCGCAAGTCGATGCTGTAATACTGCACCA	600
Qy	829	GGAATATAGCATGTGATGTGTCACGCGAGCATCTTGATCGCATGAAAAACAGTTGATCTGA	888
Db	601	GGAATATAGCAATGTGATGTGTCACGCGAGCATCTTGATCGCATGAAAAACAGTTGATCTGA	660
Qy	889	TGCATATATGGGGCAC-TCGACACAGAAATGCATGTGACCAACCCGC	934
Db	661	TGCATATATGGGGCACCTTCACACAGAAATGCATGTGACCAACCTTC	707
RESULT 11			
BM837838			
LOCUS			
DEFINITION	BM837838	691 bp	mRNA linear EST 06-MAR-2002
ACCESSION	K-EST011387	S9SNW601	Homo sapiens CDNA clone S9SNW601-76-A11 5'
VERSION	BM837838		mRNA sequence.
KEYWORDS	BM837838.1	GI:19194247	
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 691) Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.T., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S. 21C Frontier Korean EST Project 2001 Unpublished (2002) Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel.: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsungemall.kr@db.re.kr Plate: 76 row: A column: 11 High quality sequence stop: 691. Location/Qualifiers		
TITLE	1. 691		
JOURNAL	/organism="Homo sapiens"		
COMMENT	/db_xref="taxon:9606"		
	/clone="S9SNW601-76-A11"		
	/clone_11b="S9SNW601"		
	/sex="M"		
	/tissue_type="Ascites"		
	/cell_type="Epithelial"		
	/cell_line="SNW-601"		
	/lab_host="TOP10F"		
	/note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI; Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including SfiI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized with Superscript II using SfiI oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII-digested pME18S-PL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli TOP10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."		
BASE COUNT	175 a	151 c	194 g 171 t
Query Match	53.1%	Score 687.8;	DB 14; Length 691;

RESULT 13	956 bp	mRNA	linear	EST 12-JAN-2001
LOCUS BF792750				
DEFINITION 602253481F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4345629 5',				
ACCESSION BF792750				
VERSION BF792750.1				
KEYWORDS GI:12097804				
SOURCE human.				
ORGANISM Homo sapiens				
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
AUTHORS 1 (bases 1 to 956)				
TITLE NIH-MGC http://mgs.nci.nih.gov/.				
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)				
COMMENT Contact: Robert Strausberg, Ph.D. Email: c9apbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: ILAM966 row: a column: 22 High quality sequence stop: 671. Location/Qualifiers 1..956 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4345629" /clone_1db="NIH_MGC_84" /clone_2db="NIH_MGC_84" /clone_type="adrenal cortex carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: adrenal gland; Vector: pCMV-Sport6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."				
FEATURES				
Source				
BASE COUNT 234 a 236 c 284 g 202 t				
ORIGIN				
Query Match 51.9%; Score 672.4; DB 12; Length 956;				
Best Local Similarity 93.1%; Pred. No. 1e-194;				
Matches 758; Conservative 0; Mismatches 51; Indels 5; Gaps 5				
97 ATTTCACTCAGGAAAGCTGCTCAGGAGGAGAGCCCTTGCTGCTAAATAGTGGGC 156				
1 ATTTCACTCAGGAAAGCTGCTCAGGAGGAGAGCCCTTGCTGCTAAATAGTGGGC 60.				
157 TGTACACACATACACAGCCACACAGCGGTGTGATTGAACACTCTGTGCTGGGGGCT 216				
61 TGTACACACATACACAGCCACACAGCGGTGTGATTGAACACTCTGTGCTGGGGGCT 120				
217 CAGTGGCGGTGCTGCTGCTGTAACATCTACTCAACACAGAAATAGTAGCTGACGACTG 276				
121 CAGTGGCGGTGCTGCTGCTGTAACATCTACTCAACACAGAAATAGTAGCTGACGACTG 180				
277 GCTGAGGCTGAGATTGCACTGTTGCTTGGAAGGGGCACTAGAAATGACTTCTGGTGG 336				
181 GCTGAGGCTGAGATTGCACTGTTGCTTGGAAGGGGCACTAGAAATGACTTCTGGTGG 240				
337 TGTATTGACCGCTGTGTAAACATGATGGGTGGCAGGCCAATATATCTGATATATGGG 396				
241 TGTATTGACCGCTGTGTGTAAACATGATGGGTGGCAGGCCAATATATCTGATATATGGG 300				
397 GGAAGCTTAACCAACGAGGTTTATAGAAGATATCCAAAGTGTATAGAAGATCCAGGC 456				
301 GGAAGCTTAACCAACGAGGTTTATAGAAGATATCCAAAGTGTATAGAAGATCCAGGC 360				

Oy	457	ATTGGGAAGAGGAGTGACTGGTGTTCACAGGCGTATACAGCTCTCCAAAGCTGGGAG	516
Db	361	ATTGGGAAGAGGAGTGACTGGTGTTCACAGGCGTATACAGCTCTCCAAAGCTGGGAG	420
Oy	517	CTCTGTGTTCCGGCCATGAGCTCAATGATTTCTGTACCAACAGAGTTGGATTAACCTG	576
Db	421	CTCTGTGTTCCGGCCATGAGCTCAATGATTTCTGTACCAACAGAGTTGGATTAACCTG	480
Oy	577	TACTCTCTCCCGAGATTCATTTTGGATATGGCCGAGAGACACAGATGTGATTTGGT	636
Db	481	TACTCTCTCCCGAGATTCATTTTGGATATGGCCGAGAGACACAGATGTGATTTGGT	540
Oy	637	GGGAAA -CAAGGCGGTGTGCGTATAGTGAGGTA -GGCAAGGAGCTGTGTCGTC	694
Db	541	GGGAAACCAAGTGTGTGTGTGTGTGTATAGTATAGGTCAGGGAGGCTGTGTCGTC	600
Oy	695	TCAAAAGCTTGGAGCAATTTGCTACTATTTACCGAAATGACCCCATCTGTCTGTGGAG	754
Db	601	TCAAAAGCTTGGAGCAATTTGCTACTATTTACCGAAATGCA -CCCATCTGTCTGTGGAG	658
Oy	755	CTGTGATGGATGGGTCTACAGGT -GGTAAAGCTAAATGAAGTATCCGGCAATGATGTC	813
Db	659	CTGTGATGGATGGGTCTACAGGCGGTACAGCCTACTGCAAGTCAACCGGACGTGAAAGCC	718
Oy	814	GTAATTAAGTTCGACAGGAATTAAGATGTAGTGAACAGGGAGCACTTGGATGCGATGAA	873
Db	719	GTAATTAAGTTCGACAGGGAAACACATGTGCGGGCCACAGGCAACACAGGACCGGTGTA	778
Oy	874	AACAGTGTATGTATGCAATAATAGGAGCCACTCTCA	907
Db	779	AACAGTGTATGCAACCAATATGGGGCCACTCTCA	812


```
Db 601 AACTCTAATGCAACCGAGGGGCGATACAAAGGATGTGTACTTGCCTTCTTAAGAA 660
OY 1168 ATGATGAATACGTTG--CCAGCTTGATCTGCCAT-CATTGATGCCCACTTACAGAG 1224
Db 661 ATGATGAATACGTTGCGCAGCTTTGCATCTGCCATTCATGTGATGGCCACCTTACAGAG 720
OY 1225 CTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAAAATGGCCATTCAAACTTAAT 1284
Db 721 CTGACAGATGACCAAGCAAAATATCTGGGACTCAAGAACTGGG-CATTCAAACTTACT 779
OY 1285 TATTACAGATA 1295
Db 780 TATTACAGATA 790
```

Search completed: April 21, 2003, 03:59:07
Job time : 1760.62 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 22:06:49 ; Search time 4760.08 Seconds

(without alignments)
11274.088 Million cell updates/sec

Title: US-09-782-051-1_COPY_1_1844

Sequence: 1 ggcgcggcagctcgcagct.....aacctaattaccagatcc 1844

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBank:*

1: gb_ba:*

2: gb_hg:*

3: gb_in:*

4: gb_ov:*

5: gb_ov:*

6: gb_ov:*

7: gb_ov:*

8: gb_ov:*

9: gb_ov:*

10: gb_ov:*

11: gb_ov:*

12: gb_ov:*

13: gb_ov:*

14: gb_ov:*

15: gb_ov:*

16: gb_ov:*

17: gb_ov:*

18: gb_ov:*

19: gb_ov:*

20: gb_ov:*

21: gb_ov:*

22: gb_ov:*

23: gb_ov:*

24: gb_ov:*

25: gb_ov:*

26: gb_ov:*

27: gb_ov:*

28: gb_ov:*

29: gb_ov:*

30: gb_ov:*

31: gb_ov:*

32: gb_ov:*

33: gb_ov:*

34: gb_ov:*

35: gb_ov:*

36: gb_ov:*

37: gb_ov:*

38: gb_ov:*

39: gb_ov:*

40: gb_ov:*

41: gb_ov:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1844	100.0	2563	6	AX029176	AX029176 Sequence
2	1844	100.0	2677	9	AF315687	AF315687 Homo sapi
3	1840.8	99.8	2583	9	BC010681	BC010681 Homo sapi
4	1835.4	99.5	2552	9	BC016942	BC016942 Homo sapi
5	1803.8	97.8	2526	9	BC007576	BC007576 Homo sapi
6	1547	83.9	2258	9	HS082761	HS082761 Homo sapi
7	1503.8	81.6	1772	10	BC018218	BC018218 Mus muscu
8	1469.4	79.7	2510	9	HS0800298	AL049954 Homo sapi
9	999.2	54.2	2030	9	BC008349	BC008349 Homo sapi
10	999.2	54.2	2030	9	BC024325	BC024325 Homo sapi
11	995.6	54.0	5025	9	AB020635	AB020635 Homo sapi
12	979.8	53.1	1987	9	AK025372	AK025372 Homo sapi
13	651.8	35.3	1792	3	AY113501	AY113501 Drosophila
14	644	34.9	1369	9	AF035319	AF035319 Homo sapi
15	599	32.5	1333	9	BC003631	BC003631 Homo sapi
16	486.2	26.4	40330	2	AC015387	AC015387 Drosophila
17	486.2	26.4	172372	3	AC010110	AC010110 Drosophila
18	421.6	22.9	52423	2	AC020395	AC020395 Drosophila
19	421.6	22.9	80423	3	DROABDB	L07835 Drosophila
20	421.6	22.9	175335	3	AC091636	AC091636 Drosophila
21	421.6	22.9	223098	3	AE003715	AE003715 Drosophila
22	421.6	22.9	338234	3	DM031961	U31961 Drosophila
23	412.6	22.4	2057	10	MUSSAH	L32836 Mus muscu
24	411.6	22.3	5050	3	DMBX200	X13188 Drosophila
25	411.6	22.3	2067	10	BC015304	BC015304 Mus muscu
26	411	22.3	190612	2	AC129093	AC129093 Mus muscu
27	406	22.0	1302	6	AX458351	AX458351 Sequence
28	402.4	21.8	2084	9	HUMACY2	M61832 Human S-ade
29	402.4	21.8	2175	9	BC010018	BC010018 Homo sapi
30	402.4	21.8	2183	9	BC011606	BC011606 Homo sapi
31	402.4	21.8	2355	9	AK097610	AK097610 Homo sapi
32	401.4	21.8	2029	10	RATFHH	M15185 Rat S-aden
33	399	21.6	2088	5	XLA7835	AJ007835 Xenopus L
34	394.2	21.4	2129	5	XLA7835	L35559 Xenopus lae
35	375	20.3	1603	3	AY102668	AY102668 Drosophila
36	373.8	20.3	1564	3	DMARCYGEN	X95635 D. melanog
37	359	19.5	1617	3	AF080546	AF080546 Anopheles
38	358.8	19.5	81378	9	AL772411	AL772411 Human DNA
39	358.8	19.5	183584	30	AL360171	AL360171 Human DNA
40	354	19.2	1631	3	LEISADSH	M76556 L. denovani
41	346	18.8	1315	3	DDIAHA	M1937 SLime mold
42	337	18.3	384	6	AX407699	AX407699 Sequence
43	334.4	18.1	173839	2	AC118421	AC118421 Rattus no
44	322.4	17.5	37187	8	AB004537	AB004537 Schizosac
45	322.4	17.5	43757	8	SPBC8D2	AL022072 S.pombe c

ALIGNMENTS

RESULT 1
AX029176
LOCUS
DEFINITION
Sequence 1 from Patent WO9814562.
ACCESSION
AX029176
VERSION
AX029176.1 GI:10190060
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2563)
Hart,D.N.
Enzyme having s-adenosyl-1-homocysteine hydrolase (ahcy) type
activity

Pred. No. 18 is the number of results predicted by chance to have a

QY 1741 ATCTGCAATCATTTGATGCGCCACCTTACAGAGTGAAGTGAACAGCAAAATATCTGC 1800
AF315687
LOCUS 2677 bp mRNA linear PRI 22-MAR-2002
DEFINITION Homo sapiens S-adenosylhomocysteine hydrolase-like protein mRNA,
complete cds.
ACCESSION AF315687
VERSION AF315687.1 GI:16588686
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Dekker, J.W., Budhia, S., Angel, N.Z., Cooper, B.J., Clark, G.J.,
Hart, D.N. and Kato, M.
11904675
JOURNAL Identification of an S-adenosylhomocysteine hydrolase-like
transcript induced during dendritic cell differentiation
IMMUNOGENETICS 53 (12), 993-1001 (2002)
MEDLINE
PUBMED
AUTHORS Dekker, J.W., Budhia, S., Angel, N.Z., Cooper, B.J., Clark, G.J.,
Hart, D.N.J. and Kato, M.
TITLE Dendritic Cell Research, Mater Medical
Journal Submitted (23-Oct-2000) Research Institute, Level 3, Aubigny Place, South Brisbane,
Queensland 4101, Australia
FEATURES
Source
Location/Qualifiers
1. 2677
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p21"
/clone="211(1)B"
/cell_line="Hodgkin's disease cell line 1428"
1. 114
/note="derived from 5' RACE"
369. 1961
/note="PCAL: dendritic cell expressed AHCY-like protein,
expressed in dendritic cells in blood, Langerhans cells
and dermal dendritic cells; mRNA increased during monocyte
differentiation to monocyte-derived dendritic cells in the
presence of GM-CSF and IL-4"
/codon_start=1
/product="S-adenosylhomocysteine hydrolase-like protein"
/protein_id="AA126869.1"
/db_xref="GI:16588687"
/translation="MSMPDAMPPLPGVGEELKQKEIEAEKYSFPAATYKAPKQIOF
ADMOEFTPTKTRSLRSISQSDSTISASATSDSDSEYSPREKOOTNSGS
SNECYNKIKOAEFGREIEIAEODMSLISLRKROGEEPLAGATIGCTHTACTAY
LIETICAGACQCRWACNITISQNEVAAALAGTAVPAKMKSEDDFWKICDRYAM
DGMQNMILDDGGDLTHWYKTPVFKIRIVESTYVGRLLQLSAAGLCVPAAM
NVNDSVTKQKFDNLCKRESILDGKRTDVAEGKQVVCYGEVGGCCALALAG
AIVYTEIDPICALQACMDGFEVVLKNEIVQVDVICTGKNVYTRHLDKMKSC
IVCNHSHNTEIDVTSLRPELTERVRSVDVHVPDQKRVLLAEGSLNLSGTV
PFTVSTIATQALALIELYNAPEGRYKQDYLLRKMDYVYASLHPSFDMHLELT
DDQATYGLINKNGPKPYR."

BASE COUNT 657 a 648 c 708 g 664 t
ORIGIN

Query Match 100.0%; Score 1844; DB 9; Length 2677;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1844; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCGCGGCGAGGTCGGAGCTCGAGCTGCTTCTGTGCTCTTGTGGCCACCCTGCT 60
DB 115 GGCGCGGCGAGGTCGGAGCTCGAGCTGCTTCTGTGCTCTTGTGGCCACCCTGCT 174
QY 61 GTCCGGCTGCTTGGGCTGGCGAAGCAAGGCGTGGGCGACAGCAACCTCAGAGCGA 120
DB 175 GTCCGGCTGCTTGGGCTGGCGAAGCAAGGCGTGGGCGACAGCAACCTCAGAGCGA 234
QY 121 CGCAGCTCGAGCAGAGGCGCGGCGAGAGGGTGGGCGATCCGTGTGTGAGAGGCGCGC 180
DB 235 CGCAGCTCGAGCAGAGGCGCGGCGAGAGGGTGGGCGATCCGTGTGTGAGAGGCGCGC 234
QY 181 GGGCAGGCGGCGCGCGCGCCAGAGGGGAAAGAGCGGGGCGCGGCTGACCGCTGGCC 240
DB 295 GGGCAGGCGGCGCGCGCGCCAGAGGGGAAAGAGCGGGGCGCGGCTGACCGCTGGCC 354
QY 241 GGGCGCGCGGGGAAATGATGCTAGTGCATGCGCGTGCCTGGCGGGGTGGGAGGAGC 300
DB 355 GGGCGCGCGGGGAAATGATGCTAGTGCATGCGCGTGCCTGGCGGGGTGGGAGGAGC 414
QY 301 TGAAGCAGGCGCAAGAGATCGAGAGCGCCGAGAGATCTCTTCATGCGCCAGCTACCA 360
DB 415 TGAAGCAGGCGCAAGAGATCGAGAGCGCCGAGAGATCTCTTCATGCGCCAGCTACCA 474
QY 361 AGCGCGCCAGAGCAATTCATGCTGTGATGATGATGATGATGATGATGATGATGATG 420
DB 475 AGCGCGCCAGAGCAATTCATGCTGTGATGATGATGATGATGATGATGATGATGATG 534
QY 421 CCAAACTGGCGCAAGATCTTGTGCTGCTGATGCTGATGCTGATGCTGATGCTGATG 480
DB 535 CCAAACTGGCGCAAGATCTTGTGCTGCTGATGCTGATGCTGATGCTGATGCTGATG 594
QY 481 GTTCAGCTGATCTTACACAGATAGCTGATGATGATGATGATGATGATGATGATGATG 540
DB 595 GTTCAGCTGATCTTACACAGATAGCTGATGATGATGATGATGATGATGATGATGATG 654
QY 541 AAACCACTCAAGAGGCGAGCAATTTCTGTGTGAAGAATCAACAGCAGCAATTTG 600
DB 655 AAACCACTCAAGAGGCGAGCAATTTCTGTGTGAAGAATCAACAGCAGCAATTTG 714
QY 601 GAGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 715 GAGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 774
QY 661 GTCTCAGGGGAGAGGCGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 775 GTCTCAGGGGAGAGGCGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
QY 721 CCCAGACAGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 835 CCCAGACAGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 894
QY 781 CTGTAACTACTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 840
DB 895 CTGTAACTACTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 954
QY 841 CAGTGTCCGTTGAGAGGCGAGTCAAGATCACTTCTGCTGTATGACCGCTG 900
DB 955 CAGTGTCCGTTGAGAGGCGAGTCAAGATCACTTCTGCTGTATGACCGCTG 1014
QY 901 TGAACATGATGGGTGGCGAGCGCCACATGATCTGTGATGATGATGATGATGATG 960
DB 1015 TGAACATGATGGGTGGCGAGCGCCACATGATCTGTGATGATGATGATGATGATG 1074
QY 961 GGGTTTATAGAGTATCCAAAGCTGTTTAAAGAGATCCGAGGCAATGTTGAGAGAGCG 1020
DB 1075 GGGTTTATAGAGTATCCAAAGCTGTTTAAAGAGATCCGAGGCAATGTTGAGAGAGCG 1134
QY 1021 TGACTGATGTTTACAGAGCTGTATCAGCTCCCAAGGTGGAGAGCTGTGTTCCGGCCA 1080
DB 1135 TGACTGATGTTTACAGAGCTGTATCAGCTCCCAAGGTGGAGAGCTGTGTTCCGGCCA 1194

QY	361	AGGGCCCAAGAACAAATCCAGTTGTCGTATACATGACAGAGATTCCACAAATTCGCCA	420
Db	383	AGGGCCCAAGAACAAATCCAGTTGTCGTATACATGACAGAGATTCCACAAATTCGCCA	442
QY	421	CCAAAAGCGCCGAAGATCTTTGGTCGGCTGATATCACAGTCCTCCACTGACACTACA	480
Db	443	CCAAAAGCGCCGAAGATCTTTGGTCGGCTGATATCACAGTCCTCCACTGACACTACA	502
QY	481	GTTCAGCTGATCTCTACACAGATAGCTGTGATATAGGTTTCTCCCCAGAGAAAGAC	540
Db	503	GTTCAGCTGATCTCTACACAGATAGCTGTGATATAGGTTTCTCCCCAGAGAAAGAC	562
QY	541	AAACCAACTCCAAAGGCGACAGCAATTTCTGTGTGAAGAAATCAAGCAGGAGAAATTG	600
Db	563	AAACCAACTCCAAAGGCGACAGCAATTTCTGTGTGAAGAAATCAAGCAGGAGAAATTG	622
QY	601	GACGCCGGGAGATGTAATGATTCGACAGAGACATGTCCTCTGATTTACACAGAAAC	660
Db	623	GACGCCGGGAGATGTAATGATTCGACAGAGACATGTCCTCTGATTTACACAGAAAC	682
QY	661	GTGCTCAGGGGAGAACCCCTGGCTGGTCTAAATATGAGGCTGTACACATCAACAG	720
Db	683	GTGCTCAGGGGAGAACCCCTGGCTGGTCTAAATATGAGGCTGTACACATCAACAG	742
QY	721	CCACAGACAGCGGTGTATTGAGACACTCTGTCGCCCTGGGGGGCTCAGTCCCGCTGTG	780
Db	743	CCACAGACAGCGGTGTATTGAGACACTCTGTCGCCCTGGGGGGCTCAGTCCCGCTGTG	802
QY	781	CTTGTAACATCTACTCAACTCAGATGTAAGTACCTGCAGCATCTGGCTGAAGCTGGAATTG	840
Db	803	CTTGTAACATCTACTCAACTCAGATGTAAGTACCTGCAGCATCTGGCTGAAGCTGGAATTG	862
QY	841	CAGTGTTCGCTTGGAAAGGGGAGCGAAGAAATGACTCTTGAGGCTGATATGACCCCTGG	900
Db	863	CAGTGTTCGCTTGGAAAGGGGAGCGAAGAAATGACTCTTGAGGCTGATATGACCCCTGG	922
QY	901	TGAACATGATGATGGGTGGCAGGCCCAACATGATCCTGGATATGAGGGAGACTTAACCCACT	960
Db	923	TGAACATGATGATGGGTGGCAGGCCCAACATGATCCTGGATATGAGGGAGACTTAACCCACT	982
QY	961	GGGTTTATAGAAAGTATCCAAACGTGTTTAAAGAAATCCGAGGCAATTGTGAAAGAGCG	1020
Db	983	GGGTTTATAGAAAGTATCCAAACGTGTTTAAAGAAATCCGAGGCAATTGTGAAAGAGCG	1042
QY	1021	TGACTGGGGTCCACAGCCTGTATCAGCTCCCAAACCTGGGAAAGCTGTGTCTCCGGCCA	1080
Db	1043	TGACTGGGGTCCACAGCCTGTATCAGCTCCCAAACCTGGGAAAGCTGTGTCTCCGGCCA	1102
QY	1081	TGAAGTCATAGATTCGTGTACCAAAACAGAAAGTTTATTAACCTGTACTCTCCGAGAT	1140
Db	1103	TGAAGTCATAGATTCGTGTACCAAAACAGAAAGTTTATTAACCTGTGTACTCTCCGAGAT	1162
QY	1141	CCATTTTGGATGGCCTGAAAGAGACACACAGATGTGATGTTTGGTGGAAACAAAGTGTGG	1200
Db	1163	CCATTTTGGATGGCCTGAAAGAGACACACAGATGTGATGTTTGGTGGAAACAAAGTGTGG	1222
QY	1201	TGCTGTGGCTATGATGGAGAGTGGCGAAGGGCTGCTGTCTCTCTCAAAGCTCTTGGACAA	1260
Db	1223	TGCTGTGGCTATGATGGAGAGTGGCGAAGGGCTGCTGTCTCTCTCAAAGCTCTTGGACAA	1282
QY	1261	TTTGTCTCATATTAACCGAAATGAGACCCCATGTCGTCTCTGAGAGCTCGCATGATGGTTCA	1320
Db	1283	TTTGTCTCATATTAACCGAAATGAGACCCCATGTCGTCTCTGAGAGCTCGCATGATGGTTCA	1342
QY	1321	GGGTGTGTAAGCTTAATGAAGTATCCGGCAAGTCGATGTGTAATTAACCTTGACACAGAA	1380
Db	1343	GGGTGTGTAAGCTTAATGAAGTATCCGGCAAGTCGATGTGTAATTAACCTTGACACAGAA	1402
QY	1381	ATAAGATGTATGATGACACGGGAGACACTTGGATGGCATGAAAAACAGTTTATCTGATGCA	1440
Db	1403	ATAAGATGTATGATGACACGGGAGACACTTGGATGGCATGAAAAACAGTTTATCTGATGCA	1462

QY	1441	ANATGGGCCACCTCCCAACACAGAAATCGATGTATACCAAGCCTCCGCACTCCGAGCTGACGT	1500
Db	1463	ANATGGGCCACCTCCCAACACAGAAATCGATGTATACCAAGCCTCCGCACTCCGAGCTGACGT	1522
QY	1501	GGAGACGATGATGTTCTCTCGATGAGACATATGTCATGTGGCCAGATAGCAAAACAGATTGTCC	1560
Db	1523	GGAGACGATGATGTTCTCTCGATGAGACATATGTCATGTGGCCAGATAGCAAAACAGATTGTCC	1582
QY	1561	TCCCTGGCAGAGGGTGCCTACTCAATTTTGAGCTGCTCCACAGTTCCCACTTTGTTCTGT	1620
Db	1583	TCCCTGGCAGAGGGTGCCTACTCAATTTTGAGCTGCTCCACAGTTCCCACTTTGTTCTGT	1642
QY	1621	CCATCACAGCCACACACACAGGCTTTGGCAGCTATAGACTCTATATATGACCCGAGGGGC	1680
Db	1643	CCATCACAGCCACACACACAGGCTTTGGCAGCTATAGAACTCTATATATGACCCGAGGGGC	1702
QY	1681	GATACACAGAGATGTGTACTGTCTCCAAAGAAATGGATGAATACGTGGCAGCTTGC	1740
Db	1703	GATACACAGAGATGTGTACTGTCTCCAAAGAAATGGATGAATACGTGGCAGCTTGC	1762
QY	1741	ATCTGCATCATTTGATGGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	1800
Db	1763	ATCTGCATCATTTGATGGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGG	1822
QY	1801	GACTCAACAAAATGGGCCATTCAAACCTAATTTATTACAGATAC	1844
Db	1823	GACTCAACAAAATGGGCCATTCAAACCTAATTTATTACAGATAC	1866

FEATURES

Source

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: IRAP Plate: 20 Row: 4 Column: 15

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

location/Qualifiers

1. 2552

/organism="Homo sapiens"

/db_xref="LocusID:10768"

/db_xref="taxon:9606"

/clone="MGC:21453 IMAGE:3450568"


```

|||||
Db      1801 CAAAATGGGCATTCACCTATATTACAGATAC 1837
RESULT 5
BC007576                                2526 bp  mRNA  linear  PRI 12-JUL-2001
LOCUS                                     MGC:15558 IMAGE:3139729, mRNA, complete cds.
DEFINITION Homo sapiens, S-adenosylhomocysteine hydrolase-like 1, clone
ACCESSION BC007576
VERSION    BC007576.1 GI:14043176
KEYWORDS   MGC.
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens.
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 2526)
AUTHORS     Strausberg, R.
TITLE       Direct Submission
JOURNAL     Submitted (10-MAY-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
            Email: cgabs.rem@nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            Info@bcsc.bc.ca
            Steven Jones, Jennifer Aaano, Ian Bosdet, Yaron Butterfield,
            Susana Chan, Readman Chiu, Chris Fell, Erin Garland, Ran Guin,
            Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Scott,
            Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
            Series: IRAL Plate: 22 Row: h Column: 21
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, similarity but not identity to protein..
            Location/Qualifiers
                1..2526
                /organism="Homo sapiens"
                /db_xref="locusid:10768"
                /db_xref="taxon:9606"
                /clone="MGC:15558 IMAGE:3139729"
                /tissue_type="Placenta, choriochorionoma"
                /clone_1lb="NIH-MGC-21"
                /lab_host="DH10B-R"
                /note="vector: pOTB7"
                /codon_start=1
                /product="S-adenosylhomocysteine hydrolase-like 1"
                /protein_id="AA07576.1"
                /db_xref="GI:14043177"
                /translation="MATVTKAPKKQIQFADKQEFKPKTKGRSISRSISQSTDS
                YSSASVYTSDDSEVPKROKQNSKSSKSNKFNKIQAFGRREIDIAODSALIS
                LKRAQGEKPLAGAKIVGCTHTIACTAVLLETLALACQKWSKCNYSQNEVFAAL
                AEAQVAFPAKGESEDFWICIDRCVMDQNMILDDGDLTHWYKKYKYPNFKKI
                RGTVESVGVHLYOLSKAGKLCVPMNDQNVTKQKPNIVCCRSIIIDGKRTD
                VMGGRQVYVCGYGVGKCKCALKALGAIYVTEIDPLCALQCMGFYVAKINEVI
                ROYDVVITGKNKVVYREHLDRKNSCIYCNNGHSTETIDVSLRTPETLWERYKQ
                VDHVIMPDGRVYVLAEGRLNLSCSPVFPFVLSITPTQALALIELYNAPEGRYKQD
                VYLLPKMKMEYVSLHLPSPDAHLTELTDOAVYLGINKNGKPEKPNYRY"
BASE COUNT 644 a 592 c 660 g 630 t
ORIGIN

```

```

Query Match      97.88; Score 1803.8; DB 9; Length 2526;
Best Local Similarity 99.98; Pred. No. 0;
Matches 1805; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 38 GTTCTCTTGTGGCCACCGTGTCCGCGTGCCTTGCGGCTGCGGCAACAGAGCGGTG 97
DB 1 GTTCTCTTGTGGCCGCGCGTGTCCGCGTGCCTTGCGGCTGCGGCAACAGAGCGGTG 60
QY 98 GGGCCACAGACCTCAGAAAGCCGACGAGCTGACGAGAGGGCGGCAAGAGGTGGCGA 157
DB 61 GGGCCACAGACCTCAGAAAGCCGACGAGCTGACGAGAGGGCGGCAAGAGGTGGCGA 120
QY 158 TCGCGTGTGGAGAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 217
DB 121 TCGCGTGTGGAGAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 218 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 277
DB 181 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 278 GCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 337
DB 241 GCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 338 CTCCTTCATGGCCACCGTCAACCAAGGCGGCCCAAGAGCAATCCAGTTGCTGATGACAT 397
DB 301 CTCCTTCATGGCCACCGTCAACCAAGGCGGCCCAAGAGCAATCCAGTTGCTGATGACAT 360
QY 398 GCAGAGAGTCCACCAATTCGCCCAACCAATTCGCCCAACCAATTCGCCCAACCAATTC 457
DB 361 GCAGAGAGTCCACCAATTCGCCCAACCAATTCGCCCAACCAATTCGCCCAACCAATTC 420
QY 458 ACACTCTCTCCACTGACAGCTCAAGTTCAGTCTGATCTTACAGATAGCTGTGATGATGA 517
DB 421 ACACTCTCTCCACTGACAGCTCAAGTTCAGTCTGATCTTACAGATAGCTGTGATGATGA 480
QY 518 GGTTCCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 577
DB 481 GGTTCCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 578 GAACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 637
DB 541 GAACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 638 TGCTCTGATTTCACTCAGAGAAAGTGTCTCAGGGGAGAGAGAGAGAGAGAGAGAGAG 697
DB 601 TGCTCTGATTTCACTCAGAGAAAGTGTCTCAGGGGAGAGAGAGAGAGAGAGAGAGAG 660
QY 698 AGTGGGCTGTACACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 757
DB 661 AGTGGGCTGTACACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 758 GGGGCGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 817
DB 721 GGGGCGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
QY 818 AGCACTGCGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 877
DB 781 AGCACTGCGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
QY 878 CTGCTGTGTATTTACCGCTGTGTGAACATGAGTGGTGGGCGAGGCAACATGATCTCTGA 937
DB 841 CTGCTGTGTATTTACCGCTGTGTGAACATGAGTGGTGGGCGAGGCAACATGATCTCTGA 900
QY 938 TGAATGGGAGACCTTAACCCAGCTGGGTTATTAAGAGATATCCAAACGCTTTAAGAGAT 997
DB 901 TGAATGGGAGACCTTAACCCAGCTGGGTTATTAAGAGATATCCAAACGCTTTAAGAGAT 960
QY 998 CCGAGGCAATTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1057
DB 961 CCGAGGCAATTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020

```


OY	568	AACGGCGTCAAGGGGAGAAACCCCTTGCGTGGTGCCTAAATATATGGGGCTGTACACATCA	717
Db	361	AACGGCTCAAGGGGAGAAACCCCTTGCGTGGTGCCTAAATATATGGGGCTGTACACATCA	420
OY	718	CAGCCACAGACAGCGGTGTGATTGAGACACACTCTGTCCCTGGGGGCGCTCAGTCCCGCTGT	777
Db	421	CAGCCACAGACAGCGGTGTGATTGAGACACACTCTGTCCCTGGGGGCGCTCAGTCCCGCTGT	480
OY	778	CTGCTTGTAAACATCTACTCACTCAGATGAAGTACCTGCAGCAGCTGCGTAGGCTGGAG	837
Db	481	CTGCTTGTAAACATCTACTCACTCAGATGAAGTACCTGCAGCAGCTGCGTAGGCTGGAG	540
OY	838	TTGAGAGTTCGCGTTGGGAAGGGGAGTCAGGAAGATACCTTGGTGGTCTATATGACCGCT	897
Db	541	TTGAGAGTTCGCGTTGGGAAGGGGAGTCAGGAAGATACCTTGGTGGTCTATATGACCGCT	600
OY	898	GTCGAAACATGATAGGTGTGGCAGGCCCAACATGATCTCGATATATGGGGAGACTTAACCC	957
Db	601	GTCGAAACATGATAGGTGTGGCAGGCCCAACATGATCTCGATATATGGGGAGACTTAACCC	660
OY	958	ACTGGGTTTATAGACGTAATCCAAACGTGTTTAAAGAATCCGAGGCATTGTGAAAGAGA	1017
Db	661	ACTGGGTTTATAGACGTAATCCAAACGTGTTTAAAGAATCCGAGGCATTGTGAAAGAGA	720
OY	1018	GGGAGCATGGGGTTCACAGGCTGTATCAGCTGCCAAACCTGGGAAGCGCTGTGTTCCGG	1077
Db	721	GGGAGCATGGGGTTCACAGGCTGTATCAGCTGCCAAACCTGGGAAGCGCTGTGTTCCGG	780
OY	1078	CCATGAACGTCATGATTTCTGTTACCAAAACAGAACTTATATACTGTACTCTCCGAG	1137
Db	781	CCATGAACGTCATGATTTCTGTTACCAAAACAGAACTTATATACTGTACTCTCCGAG	840
OY	1138	AATCCATTTTGGATGGCCTTAAGAGACACAGATGTATGTTTGGTGGGAAACAAGTGG	1197
Db	841	AATCCATTTTGGATGGCCTTAAGAGACACAGATGTATGTTTGGTGGGAAACAAGTGG	900
OY	1198	TGGGTGTGGCTATGGGAGGTAGGGAAGGGGCGTGTCTCTCTCACAAGCTCTTGGAG	1257
Db	901	TGGGTGTGGCTATGGGAGGTAGGGAAGGGGCGTGTCTCTCTCACAAGCTCTTGGAG	960
OY	1258	CAATTGTCTACATTTACCGAATAGACCCCATCTGTCTCTCGAGGCGCTGCATGATGGGT	1317
Db	961	CAATTGTCTACATTTACCGAATAGACCCCATCTGTCTCTCGAGGCGCTGCATGATGGGT	1020
OY	1318	TCAGGGTGTGAAAGCTAAATGAAGTCATCCGGCACTGATGTCTGTAATTACTGCACAG	1377
Db	1021	TCAGGGTGTGAAAGCTAAATGAAGTCATCCGGCACTGATGTCTGTAATTACTGCACAG	1080
OY	1378	GAAATTAAGAAATGTAGTGCACACGGGAGACACTTGGATGCGATGAAAAACAAGTTATACGAT	1437
Db	1081	GAAATTAAGAAATGTAGTGCACACGGGAGACACTTGGATGCGATGAAAAACAAGTTATACGAT	1140
OY	1438	GCAATATGSGGCACACTCCACAACAGAAATGGAATGTGACAGCCTCCGCACTCCGGAGCTGA	1497
Db	1141	GCAATATGSGGCACACTCCACAACAGAAATGGAATGTGACAGCCTCCGCACTCCGGAGCTGA	1200
OY	1498	CGTGGAGCGAGTACGTTCTCAGGTGAGACATGTCTCTGGCAGATGGAACAACAGATTG	1557
Db	1201	CGTGGAGCGAGTACGTTCTCAGGTGAGACATGTCTCTGGCAGATGGAACAACAGATTG	1266
OY	1558	TCGTCGTGGCAGAGGGTCTGTCTCAATAATTTGAGCTGTCCACAGTTCGCCACTTGTTC	1617
Db	1261	TCGTCGTGGCAGAGGGTCTGTCTCAATAATTTGAGCTGTCCACAGTTCGCCACTTGTTC	1320
OY	1618	TGTGCATACACAGCCACAACAACAGGCTTGGCACTGATATACTCTAATAATGACACCCGAG	1677
Db	1321	TGTGCATACACAGCCACAACAACAGGCTTGGCACTGATATACTCTAATAATGACACCCGAG	1380
OY	1678	GGCAGATCAAGACAGAGATGTACTCTTGCTTCCTTAAGAAATGGATGAATACGTTGGCAGT	1737
Db	1381	GGCAGATCAAGACAGAGATGTACTCTTGCTTCCTTAAGAAATGGATGAATACGTTGGCAGT	1440
OY	1738	TGCATCTGCCATCATTTGATGCCACCTTACAGAGCTGACAGATGACCAAGCAATAATTC	1797

Db	1441	TGCACTGCGCATCTTTGGATGCCACCTTACAGAGCTGACAGTGCACAGGCCAAGAAATATTC	1500
Qy	1798	TGGGACTCAACAAAAATGGGCCATTCAAACCTAATATTATACAGATAC	1844
Db	1501	TGGGACTCAACAAAAATGGGCCATTCAAACCTAATATTATACAGATAC	1547

RESULT	7
LOCUS	BC018218
DEFINITION	BC018218 1772 bp mRNA linear ROD 07-AUG-2002
ACCESSION	Mus musculus, S-adenosylhomocysteine hydrolase-like 1, clone MGC:18748 IMAGE:4007102, mRNA, complete cds.
VERSION	BC018218
KEYWORDS	BC018218.1 GI:17390492
SOURCE	MGC.
ORGANISM	house mouse. Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
1 (pages 1 to 1772)
Strausberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian

REMARK	COMMENT
NIH-MGC Project URL: http://mgc.ncl.nih.gov	Contact: MGC help desk

Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNI)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarone, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M., Yoon, V.S., Kovys, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNl at: http://image.llnl.gov/Series1:IRAK_plate.23.Row:n.Column:6. This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction, Similarly but not identically to protein.

FEATURES	Location/Qualifiers
source	1. .1772

CDS

/organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="CZECH 11"
 /clone="MGC:16748 IMAGE:4007102"
 /tissue="type="Mammary tumor metastasized to lung."
 MMTV-LTR/Mtl1 model. Expression driven by an MMTV-LTR
 enhancer."
 /clone.lib="NCI CGAP Lu30"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 191. 1693
 /codon_start=1
 /product="5-adenosylhomocysteine hydrolase-like 1"
 /protein_id="AAH18218.1"
 /db_xref="gi:17390493"
 /translation="MAVTVYAKKQIOFADDDMEFTKPTKGRSLRSISQSSDPS
 YSSASVYSDSDSDSPREKQSSNGSSNCFYKNKQAFGRREITLADDMALSL
 LKRRAGSGKPLAGKITYGCHITNQRVLVLEETCALGACGRSACNITYSTONELVAAAL
 LEAGVAVFAMKGSSEDDFMWIDICVAMDMQANMLLDGGGLTWTYKTKIPNFKKI
 KQIVESYTVGHRTYLSLKSGKLCVPMANVNSVTKQKDNLYCCRESITDGLKRTVEY
 WFKGQVYVGVGVGEVSKGCALAKGALVYITETIDPICALQAMDFEFTVRLKNEYI

RESULT 9	
LOCUS	BC008349
DEFINITION	2030 bp mRNA linear PRI 12-JUL-2001
ACCESSION	BC008349
VERSION	1
KEYWORDS	clone IMAGE:3536052, mRNA, partial cds.
SOURCE	BC008349.1 GI:14249935
ORGANISM	Homo sapiens.
REFERENCE	Homo sapiens Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2030) Strausberg, R. Direct Submission Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
AUTHORS	NIH-MGC Project URL: http://mgc.nci.nih.gov
TITLE	Contact: MGC help desk
JOURNAL	Email: cgabs-remail.nih.gov Tissue Procurement: DCTD/DRP CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca
REMARK	Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Nan Gulin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lies Prabhu, Parvaneh Saeedi, Jacqueline Schell, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsel, Netaja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IMAGE Plate: 23 Row: 5 Column: 18. Location/Qualifiers 1..2030 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3536052" /tissue_type="Lung, small cell carcinoma" /clone_lib="NIH.MGC.7" /lab_host="DH10B-R" /note="Vector: pOTB7" <1..1838 /codon_start=3 /product="similar to S-adenosylhomocysteine hydrolase-like 1." /protein_id="AA08349.1" /db_xref="GI:14249936" /translation="MSYQVVSAAARVPEVELEKDLSPSEASOLGISTAVGAMAPP AGGDEPAPAPAEERPPVPGSGSPAPALSPAAGVPOASMKRSDPHORRDGSG EALVSPGTEVARETRKQIOFADQENFKRPTKIGRGLSRISQSSSTDSVSAAR SYTSPDSTSPBDKOOKNGSGSDCVKNTKQAEPRERTEIEDECMPLMARRRA OGERPLAGAKIVCTGTHITPAQTVAMETLGALGACRRAACITSTLNEVAAALAESEF PVFMKGESEDEFWCIDRCVNVGWPNNMILDDGSDLTHTYTKRYNMFKKKIGIVE ESVYVHRLYOLSRAGKICVPAANVNDVTKQKDNLYCCSESTLIDLRKTTDMFGG KVOVVGCGEYGGKCAALAKMGSIVYVTEIDPICALQACDGRFLVAKNEIVIQVDI KIVTGKKNVYTRHEHIDRMKNSICVCMNGHSENTEDIVASLTPELTWERYASQVDHYI MPDGRYVLAEGRLNLSCSTVPTVLSLTAATQALALETVNAPGRYKODVYLLP KMDRYVASLHPPTFDALHTLTETDQAKYLGILNNGPKPKYTYK"
CDS	BASE COUNT 539 a 492 c 548 g 451 t ORIGIN
Query Match	54.2%; Score 999.2; DB 9; Length 2030;
Best Local Similarity	75.1%; Pred. No. 1.8e-190;
Matches 1279; Conservative	0; Mismatches 413; Indels 12; Gaps 2;

OY	1220	AGGCAAGGGCGTCTGTGGTCCTCCTAAAGACGCTTGAGACAATTTGCTCATTAACGGAAT	1279
Db	1211	GGGGAAGAAGGTCGTCTGTGCTCCCTAAGAACCCATTGGGCTCATTTGTATTAACGAAT	1270
OY	1280	GCACCACATCTGTGCTGTCGACAGGCCCTGCATGGATGGGTTACAGGTCGTAAAGCTTAATGA	1339
Db	1271	TGACCCCATCTGTGCTGCTGACAGCCTGTATGGATGGATTTGCACCTGGTGAATTAAATGA	1330
OY	1340	ACTCATCCGGCACAGTCGATGTCGTATAACTTGCACAGAAAATAAGAAATGTAGTACACG	1399
Db	1331	GGTCATCCGACAAAGTAGACATGTTATTAACCTGTACAGAGTAAACAAGAAATGTGTAAACG	1390
OY	1400	GGAGCATTTGGATGCCATGAAAAACAGTTGTATCTATATGCATAATATGGCCACTCCAACAC	1459
Db	1391	AGAGCACTTGGACCCCTATGAAGAAATAGCTGCATCTTTGTAACTAATGGGACATTCACACAC	1450
OY	1460	AGAAATCGATGTCGACAGCTCCGACACCTCCGAGCTGACGTCGGAGGCGAGTACGTTCA	1519
Db	1451	AGAGATTCGACGGGCGAGCTCTGGACACACAGAACTGACCTGGGAGCGAGTAAATCTCA	1510
OY	1520	GGTGACCATGTCATCTGTGGCCAGATGGGCAAAAGATTTCCCTCCAGAGAGAGTCTCT	1579
Db	1511	AGTTACCATCATGATATGATGGCCGTGATGGCAAGAGATATGACTGCTGGCAGAGGGCGCT	1570
OY	1580	ACTCAATTTGAGCTGCTCCACAGTCCACCTTTGTTCTGTCATACAGCCACAACACA	1639
Db	1571	GCTGAACCTTGAAGCTCTCCACAGTCCCTAATTTGTGCTCTCAATCACAGTACTACACA	1630
OY	1640	GGCCTTGGCACATGATAGAACTCTATATAATGCAACCCAGGGGCGATACAAAGAGATGTGA	1699
Db	1631	GGCTTTGCTCTTGAATGAGCTTTAACAATGCTCTAGAGGTGCTCTTAAGCAGAGATGTGA	1690
OY	1700	CTTGCTCTCTAAGAAATAGGATGATATACGTTGCCAGCTTGATCTGCATCATTTGTATGAC	1759
Db	1691	CCTGTGGCCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1750
OY	1760	CCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAAATGGGCC	1819
Db	1751	CCACTTGCACAGAGCTGACAGATGAAACAGGCCCAAGATCTGGGACTCAACAAAATGGGCC	1810
OY	1820	AFTCAACCTAATTTATTAACAGATA	1843
Db	1811	CTTCAAGCTAATTAATTAACAGATA	1834
RESULT 10			
LOCUS	BC024325	5052 bp	mRNA linear PRI 12-MAR-2002
DEFINITION	Homo sapiens, KIA00828 protein, clone MGC:21525 IMAGE:3907552,		
ACCESSION	BC024325		mRNA, complete cds.
VERSION	BC024325.1	GI:19353091	
KEYWORDS	MGC.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 5052)		
AUTHORS	Strausberg,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-MAR-2002) National Institutes of Health, Mammalian		
	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov		
COMMENT	Contact: MGC help desk		
	Email: cgapbs-ref@mail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)		
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome		
	Center, Stanford University School of Medicine, Stanford, CA 94305		
	Web site: http://www-shgc.stanford.edu		

Contact: (Dickson, Mark) mcdpax11.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LML at: <http://image.lml.gov>
 Series: IRAC Plate: 22 Row: 1 Column: 22
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES
 source
 Location/Qualifiers
 1..5052
 /organism="Homo sapiens"
 /db_xref="locusID:23382"
 /db_xref="taxon:9606"
 /clone="MGC:21525 IMAGE:3907552"
 /tissue_type="uterus" leiomyosarcoma"
 /clone_1lb="NH_MGC_71"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 55..1890
 /codon_start=1
 /product="KIAA0828 protein"
 /protein_id="AAH24325.1"
 /db_xref="GI:19353092"
 /translation="MSVGVVSAALAAKVEVELKDLSPSEASOLGSLTAIVGAMAPPAGGDEAPAPARPPVPVGGSGPAAALSPAACKVPOASMRSDGHHORHBDGGALVSPDGYTPAERYTKQIQAPDQKQEPKRTKIGRSLSSIOSSDISSAAASTYSDSDETSFRDQKNSKSSDFCVKNLQKQEGRELEIQEHPALMAALKRRAQEPPLAGAKIVGCTHTTAQAVLMETLGAAGCARNACIVYSTLEVAALASGPFPVFMKGGSEDEDFWMCIDRCVNVGMSPTMMLDGGDLTHVIRKYINMEFKIGIVEEBSVYGHSLYOLSKAGKICVPAMVNVSPYKQKPDNLYCSESLLDGIKRTMDPFGGKOVVYGGGEVYEGKCCAAKAMGSIVTVELDICALQACDGRALKVNEITVQNDIIVITCTGKNVYTRPHLDLBMKRSCLVCMNHSHTSIDVAASLTPPELTERNSOVDHYIWPDKRKYLLAEGRLLNLSCTVFFVFLSSTATIQAALILEYVAPGRKQDYLLPKMDEYVASLPLPFDAHLTELTEDEQAKYGLNLNKGPFKPYRYR"

BASE COUNT 1300 a 1184 c 1145 g 1423 t
 ORIGIN

Query Match	Best Local Similarity	54.2%	Score 999.2;	DB 9;	Length 5052;
Matches 1279;	Conservative 0;	Mismatches 413;	Indels 12;	Gaps 2;	
Oy 149	GATGGCGATGCGCTGTGTGAGAGGGCGCGCGGGGCGGCGCCAGAGGGGGA	208			
Db 186	GGCGGGCGGTGGAGAGCCTGAGGCTTCAGCTCCCGCGGAGCGGCCCGCGTCCCGG	245			
Oy 209	AAGAGGCGGGGCGGCGGCTGACCGGCTGGCCCGGGCGGGGGAATGTGATGCTGA	268			
Db 246	CCCGGGCTCGGGGCGCGCGCGCTCTGTCAGCCCGCGCGGGAAGGTGCTCAGGGCTG	305			
Oy 269	CGCATGCGCGCTGCCCGGGGTGGGGGAGAGAGCTGAAGCAGGCCAAGAGGATCGAGGAGC	328			
Db 306	GGCCCTTGAAGGGAGAGCGAGCCACATCACAGACACAGCGGAGCCGCGAGCGGGGAGGC	365			
Oy 329	CGAGAAGTACTCTTCATGAGCCACCGTCACCAAGGCGCC-----CAAGAACAAT	379			
Db 366	CCTGCTGAGCCCCGAC---GGCACCGTACACGAGGCGCGCGCACAGTCAAGAACCAAT	422			
Oy 380	CCAGTTGCTGATGACATGACAGAGATTACCAAAATTTCCCAACAAACTGGCCGGAATG	439			
Db 423	CCAGTTGCTGACCAAGAAAGAAATTCACAAAGAGTCCCAACAAAAATTGACGTCGCTC	482			
Oy 440	TTTGCTGCTGCTGATCTACAGTCTCTCCACATGACAGCTACAGTTAGTGCATCTTAC	499			
Db 483	TTTGCTGCTGCTTCCATTTTCTCAGTCTATCTATGTACAGCTACAGCTACAGGGGCTTAC	542			
Oy 500	AGATAGCTGTATGATGAGGTTTCTCCCGGAGAGAGAGCAACCACTCCAAAGGCGAG	559			
Db 543	AGATAGCTGTATGATGAGATGAGCATGCCCAAGGAGACAGCAAAAAGAACTTAAGGAG	602			
Oy 560	CAGCAATTTCTGTGTGAAGAACATCAAGCAGGAGCAAAATTTGGAGCGCCGGGATTTGACAT	619			

Db 603 CAGTGAATTCTGTTAGAACATCAACAGGAGATTGGACGAGAGAAATGAAAT 662
 QY 620 TCGAGAGCAAGACATCTGCTGTGATTCTACAGAGAAAGCTCTCAGGGGAGAAACC 679
 Db 663 TCGTGAAGCAAGAAATGCTGATGATGCTTGGAGAGAGAGAGAGAGAGAGAGAGAG 722
 QY 680 CTGGGTGGTGGTAAATAGTGGGCTGACACATCAAGCCGACAGAGGCTGTGAT 739
 Db 723 TTGGGTGGGAG 782
 QY 740 TGAAGACATCTGCTGCTGGGGGCTCAGTCCGCTGCTGTGTAACATCTACAC 799
 Db 783 GGAAGCTGTGGTGTCTGGGGGCTCAGTCCGATGGGCTGCTGCAACATCTATCCAC 842
 QY 800 TCGAATGAGATGCTGACAGACTGGCTGAGGCTGAGAGTGGCTGTGCTTGAAGG 859
 Db 843 TCTCAATGAGATGCTGCTGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 902
 QY 860 CGAGTGAAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 919
 Db 903 AGAGTGAAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962
 QY 920 GGGCAACATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979
 Db 963 GCGAAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
 QY 980 AAGCGTGTAAAGATCCGAGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1039
 Db 1023 CAACATGTTAAAGAAATCAAGGGCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1082
 QY 1040 GTATCACTCTCCAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1099
 Db 1083 GTACCAAGCTGCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142
 QY 1100 TACCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1159
 Db 1143 CACCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1202
 QY 1160 GAGGACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1219
 Db 1203 AAGGACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1262
 QY 1220 AGGCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1279
 Db 1263 GGGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322
 QY 1280 CGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1339
 Db 1323 TGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1382
 QY 1340 AGTCAATCCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1399
 Db 1383 GGTCAATCCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1442
 QY 1400 GAGCACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1459
 Db 1443 AGAGCACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1502
 QY 1460 AGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1519
 Db 1503 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1562
 QY 1520 GGTGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1579
 Db 1563 AGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1622
 QY 1580 ACTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1639
 Db 1623 GGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1682
 QY 1640 GGTGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1699
 Db 1683 GGTGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1742

QY 1700 CTGCTCTCTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1759
 Db 1743 CCGTGGTCCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1802
 QY 1760 CCAGCTTACAGAGCTGACAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1819
 Db 1803 CCAGCTTACAGAGCTGACAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1862
 QY 1820 ATTCAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1843
 Db 1863 CTCAAGCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1886

RESULT 11

AB020635 5025 bp mRNA linear PRI 16-JUN-1999
 Locus Homo sapiens mRNA for KIA0828 protein, partial cds.
 DEFINITION
 AB020635
 ACCESSION
 VERSION
 AB020635.1 GI:4240144
 KEYWORDS
 SOURCE Homo sapiens adult male brain cDNA to mRNA, clone_11b:pb1uescript11
 SK plus clone:h04230.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (sites) Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Hirose,M., Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O. Prediction of the coding sequences of unidentified human genes. xii. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
 DNA Res. 5 (6), 355-364 (1998)

JOURNAL MEDLINE 2 (bases 1 to 5025)
 REFERENCE 99156230
 AUTHORS Ohara,O., Suyama,M., Kikuno,R., Nagase,T. and Ishikawa,K.
 TITLE Direct Submission
 JOURNAL Submitted (02-DEC-1998) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology, Yana 1532-3, Kisarazu, Chiba 292-0812, Japan (E-mail:cdna1nfo@kazusa.or.jp, Tel:+81-438-52-3913, Fax:+81-438-52-3914)

FEATURES

source location/Qualifiers
 1..5025
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="h04230"
 /sex="male"
 /tissue_type="brain"
 /clone_11b="pb1uescript11 SK plus"
 /dev_stage="adult"
 1..5025
 /gene="KIA0828"
 <1..1860
 /gene="KIA0828"
 /codon_start=1
 /product="KIA0828 protein"
 /protein_1d="BA74651.1"
 /db_xref="GI:4240145"
 /translation="EPVVAAEVMSVQVYSAARAKVPEVLEKDLSPSEASQGLST
 AAVGAPAPGAGDPEAPAPAEERPVPGSPAPALSPAAKRVQASAMKRSDDPH
 OHORHHDGGEALVSPDGTVEAPRYTKITOPADOKGEFNRKPKIGRSLRSISOS
 TBSYSRAATGSDDETSPRDKQKNSGSSDFCKNKIKQAFGRRELIQAQBPAP
 LNAIKRRAQGERPLAGAKIVGCTHTTAQAVLMEIAGALQACRMACNTYSTLNEA
 AALAESGEPVFAKGSSEDDFWWCIDRCVNEBGMOMLDDGDLTHNYKYPNF
 KIKIGVESVTVGHRLYOLSKAGKLCVAMVNDVTRKQFENLCCESISIDGDKR
 TDDMGKGVQVYGVGEGKCCALKNMGSIVYVEIDPICALQADGDFLYLN
 EYIRVDYITGCKNKNVYTRHEDLKNKSCVIVNGHSTETDVAISLRPELTWRV
 RSOVDVIMPDKRIVLALBGRILNLSCTSPVFFVSTRITTDALALIELYNABERY
 KODVYLLPKMDEYVASLHPTFDALHTELDQAKYLGINKGPKPYIKY"

BASE COUNT

ORIGIN 1281 a 1179 c 1136 g 1429 t
 Query Match 54.0%; Score 995.6; DB 9; Length 5025;

Best Local Similarity 74.6%; Pred. No. 9, le-190:
Matches 1266; Conservative 0; Mismatches 429; Indels 3; Gaps 1;

149 GGTGGCGATCGGCTCGAGAGGCGCCGCGGCGAGCGGCGGCGAGAGGGGGA 208
159 GGGGGGGGGGAGACCTGAGGCTCCAGCTCCGCGGAGAGGGGCCCCGGTCCCGG 218
209 AAGAGCGGGGGGGGGGCTGAGCCCTGCGCGGCGGCGGAGAGTTCATGCTTGA 268
219 CCGGGGCTCGGGGCGCGCGCTCTCAGCCCGCGGAGAGTTCCTGAGCGCTG 278
269 CGGATCGCGCTCGCGGGGGTGGGGAGGAGTGAAGCGAGGCGCAAGAGATTCAGAGCGC 328
279 GGCATGAAAGGAGGAGACCCATCATCCAGACGAGCGGCGGCGGCGGCGAGG 338
329 C---GAGAGTACTCTTCATGCGCACCGTCAACAGGCGCGGCGGAGAGCAATTCAGTT 385
339 CCGGTCAGACCGCGGAGGAGACCGTCAACGAGGGCGGCGGAGTCAAGAGATTCAGTT 398
386 TCGCTGATGACAGGAGAGTTCACCAATTCGCCACCAAAATGCGCGAGATCTTGTTC 445
399 TCGTACCGAGAGCAAGATTCACCAAAATGCGCGAGATTCAGTCTCTTGTTC 458
446 TCGCTGATGATCAGTCTCCAGTACAGAGTTCAGTCTGAGTCTTACAGAGTAG 505
459 TCGTTCATTTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCT 518
506 CTTCTGATGATGAGTCTCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 565
519 CTTCTGATGATGAGAGATTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 578
566 TTTCTGTGAGAGACATCAG 625
579 CTTCTGTGAGAGAGATTCAG 638
626 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 685
639 ACAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 698
686 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 745
699 TGGAGCGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 758
746 ACTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 805
759 TCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 818
806 TGAAGATG 865
819 TGAAGATG 878
866 ACAAGATG 925
879 AAGAGATG 938
926 CATG 985
939 CATG 998
986 GTTTAAGAGATG 1045
999 GTTTAAGAGATG 1058
1046 GGTCTCAAGATG 1105
1059 ACTGTCAAGATG 1118
1106 ACAAGATG 1165
1119 ACAAGATG 1178
1166 CACAGATG 1225

Db 1179 AACAGACATG 1238
Qy 1226 GGGCTGT 1285
Db 1239 AGGATG 1298
Qy 1286 CATCTGT 1345
Db 1299 CATCTGT 1358
Qy 1346 CCGGCAAGT 1405
Db 1359 CCGGCAAGT 1418
Qy 1406 CTTGATG 1465
Db 1419 CTTGATG 1478
Qy 1466 CGATGTGATG 1525
Db 1479 TGAAGTGTGATG 1538
Qy 1526 CCATGTGATG 1585
Db 1539 CCATGTGATG 1598
Qy 1586 TTTGATG 1645
Db 1599 CTTGATG 1658
Qy 1646 GGCATGTGATG 1705
Db 1659 TGCCTGTGATG 1718
Qy 1706 TCCATG 1765
Db 1719 GCCCAAGATG 1778
Qy 1766 TACAGATG 1825
Db 1779 GACAGATG 1838
Qy 1826 ACCTAATG 1883
Db 1839 GCCTAATG 1896

RESULT 12
AK025372 1987 bp mRNA 1linear PRI 29-SEP-2000
LOCUS AK025372
DEFINITION Homo sapiens CDNA: FLJ21719 fls, clone COLF0094.
ACCESSION AK025372
VERSION AK025372.1 GI:10437875
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens colon mucosa CDNA to mRNA, clone_1lb:COLF.
clone:COLF0094.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (sites)
REFERENCE
Watanabe, K., Kumagai, A., Itakura, S., Yamazaki, M., Tashiro, H.,
Ota, T., Suzuki, Y., Odayashi, M., Nishitani, T., Tanaka, T.,
Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project
2 (bases 1 to 1987)
TITLE JOURNAL
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Odayashi, M., Nishitani, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286).

COMMENT

Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcriptase errors of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our web site (<http://fruitfly.berkeley.edu>) or send email to cdna@fruitfly.berkeley.edu.

FEATURES

source

Location/Qualifiers
1..1792
/organism="Drosophila melanogaster"
/strain="y: cn bw sp"
/db_xref="taxon:7227"
/map="89E10-89E10"
1..1792

gene

/gene="hcy89E"
/note="alignment with genomic scaffold AE003715"
/db_xref="FLYBASE:FBgn0015011"
221..1699

CDS

/gene="hcy89E"
/note="Longest ORF"
/codon_start=1
/product="RE58316p"
/protein_id="AA029506.1"
/db_xref="GI:21064553"
/db_xref="FLYBASE:FBgn0015011"
/translation="MAKPEETFDLSADKTAVKSSIEARFSDVSTCSSTCET
GSSDEEDVSPKDNHNSAGGDFPCVSKSAPFREFIESEMPGIMTRAKD
EKPLKAGIVCTHYNAOSAVILFVLOLGAIVMAACITSTONVAALAEAGPI
FAMGEFEFEFMMCLDRAIVSDGQPNILIDGGATHTLMTKDYDPAFAGIYES
VTGHRILMTSKGSLTPALIVNDVSRKNDPTTYTCDLSLKRITDMEGKO
VTCIGDVGKCAQSLKGGCTIVTVEIDPICALAANDGRVVRVRLNVDIVV
TATGNKNTVTRDHNMRKNGCLTWNMGSCSEIDVNGLTPELTVERNSVDHTRMP
DGRMILAEGLVNLSCSTISSFYVASSTQALATLELFAFAPRYKSDVYLPRKM
DEVASLHAEFDALHLELTDGOSKFMCLNKGAPKANYRY"

BASE COUNT 463 a 425 c 483 g 421 t
ORIGIN

Query Match

Best Local Similarity 35.3%; Score 651.8; DB 3; Length 1792;
67.1%; Pred. No. 1e-120;
Matches 939; Conservative 0; Mismatches 457; Indels 3; Gaps 1;

QY 445 CTCGCTCATCTCAGAGTCTCCAGTCTGACAGTCTAGTCTGATCTCTACAGATA 504
DB 300 CTCGCGGGTTTCTGATGTGTCACCTGCTCATCTGCTTCCACCGGAGCT 359
QY 505 GCTGATGATGAGGTTTCTCCCGAGAGAGAGCAACCACTCCAGGGGAGAGCA 564
DB 360 CCGAGAGAGAAAGCTCTCCCGAGAGAGACATCAAGCGCACTCCCGGAGGACT 419
QY 565 ATTCTGTGTAAGACATCAAGAGGAGAAATTGGAGCCGGAGATTGAGATTGAC 624
DB 420 ACTTTTGCATGAGAGACATCTCGAAGATGATTTGGAAGGCGGAGATCGAGT 479
QY 625 ACAGAGACATGCTCTGATTTCTACTAGGAAAGCTGCTGAGGGGAGAGACCTTG 684
DB 480 AGTCGAGATGCGCGGACATCACTCTGAGGAAAGAGGAGGAGGAAACCCCTTA 539
QY 685 CTGGGTCTAAATAGTGGCTGACACATCAAGCCAGAGAGAGGAGGTTGATTGAG 744
DB 540 AAGGTGCCAATATCTGTCGATGACACCCACGTCATGTCGAGTGTCTGATGAGA 599

QY 745 CACTGTGTGCTTGGGGGCTCACTGCCCGCTGTGCTGTGTAACATCTACTCACTGAGA 804
DB 600 CCTCTGCCAAATGGGCGGCACAGATTGCTGGGGCTGCTCAACATTTATTCACACAAA 659
QY 805 ATATAAGTGTGACAGACTGTGAGCTGAGCTGAGTGTGAGTGTGAGAGGCGACT 864
DB 660 ACCCATGTGGCGGCTGTGGGAGAGCGGGAATTCGATCTTCGCTGGCGGAGAGA 719
QY 865 CAGAAATGACTCTGTGGTGTATTGACCGCTGTGTAACATGATGATGGTGGAGCGCA 924
DB 720 CGAGAGAGAGTCTGTGTGCTTGAGACAGGGCATCTCTCCAGGCTGGCAGCGCA 779
QY 925 ACATGATCTGTGATGATGGGGAGACTTAACCACTGGGTTATTAAGAAATATCAACG 984
DB 780 ACCTAATCTGAGACAGCGGGGAGATGCCACGACCTTATGTCTCAAGAAATACCCGACT 839
QY 985 TGTTTAAGAAATCCAGAGCATTTGTGGAAGAGACCGGATGTTGTTCCAGCTGATC 1044
DB 840 ACTTCAAGGCAATCGGGGCTGTGGAGAAAGTGTGACCGGGGATCCAGCGGCTGTACA 899
QY 1045 AGCTCTCAAAAGCTGGGAAAGCTGTGTGCTCCGGCATGAACTCAATGATTCTTACCA 1104
DB 900 TGCTGTCAAAAGGCGGAAACCTTACTGTTCGGGCATCAACGTTAAGACTAGTGACA 959
QY 1105 AACAGAACTTTGATTAATCTTACTGCTGCCAGAAATCCATTTTGATGGCTGAGAGCA 1164
DB 960 AGAACAAATTTGATTTCTTCTACAGCTGTGCTACCTCACTCCAGTCTCAAAAGCA 1019
QY 1165 CCACAGATGATGATTTGGGGAAACAAAGTGGTGGTGTGCTATGGAGGTAGGCA 1224
DB 1020 CCAGGATATTAATGTTTGGGGAAAGAGTGGTGTGCTGTGGTATGGAGTGGGAA 1079
QY 1225 AAGGCTGTGTGCTGTCTCTCAAAAGCTTCTTGAGACATTTCTATATACCAATGACCC 1284
DB 1080 AAGGCTGTGCGCAAGCTCCTCAAGAGGCAAGAGATGTTTATGTTACGAGAGTGGATC 1139
QY 1285 CCATCTGTGCTGTGACAGGCTGATGATGATGATGATGATGATGATGATGATGATG 1344
DB 1140 CCATATGATGCTCTCAAAAGCTGATGATGATGATGATGATGATGATGATGATGATG 1199
QY 1345 TCCGCAAGCAATGTCGATTAATGATGATGATGATGATGATGATGATGATGATGATG 1404
DB 1200 TCAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1259
QY 1405 ACTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1464
DB 1260 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1319
QY 1465 TCGATGTGACAGCTTCCGACCTCCGAGCTGACAGTGGAGAGATGATGATGATGATGATG 1524
DB 1320 TTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1379
QY 1525 AACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1584
DB 1380 ATCAATCAAGTGTGCGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1439
QY 1585 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1644
DB 1440 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1499
QY 1645 TGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1704
DB 1500 TGGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1556
QY 1705 TTGCTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1764
DB 1557 TGCCAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1616
QY 1765 TTACAGAGCTGACAGATGACCAAGCAAAATATGTGGAGTCAACAAATATGAGGCAATTA 1824
DB 1617 TCACGAGAGCTCAGAGATGACAGATGACCAAGTATGATGATGATGATGATGATGATGATGATG 1876

```

QY 1825 AACCTAATTTACAGATA 1843
DB 1677 AACCAATTAACAGATA 1695

RESULT 14
AF035319
LOCUS Homo sapiens clone 23931 mRNA, partial cds. 1369 bp. mRNA linear PRI 07-JUL-2000
DEFINITION AF035319
ACCESSION AF035319.1 GI:2661082
KEYWORDS FLI_CDNA.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Andersson,B., Wentland,M.A., Ricafrente,J.Y., Liu,W. and Gibbs,R.A.
TITLE A "double adaptor" method for improved shotgun library construction
JOURNAL Anal. Biochem. 236 (1), 107-113 (1996)
MEDLINE 96207227
PUBMED 8619474

REFERENCE
AUTHORS 2 (bases 1 to 1369)
Yu,W., Andersson,B., Worley,K.C., Muzny,D.M., Ding,Y., Liu,W.,
Ricafrente,J.Y., Wentland,M.A., Lennan,G. and Gibbs,R.A.
TITLE Large-scale concatenation cDNA sequencing
JOURNAL Genome Res. 7 (4), 353-358 (1997)
MEDLINE 9110174
PUBMED

REFERENCE
AUTHORS 3 (bases 1 to 1369)
Yu,W., Sarginson,J. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-1997) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA
FEATURES
source
location/Qualifiers
1..1369
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="23931"
/sex="female"
/lisue.type="brain"
/clone.lib="11NB"
/dev_stage="infant"
/note="This clone is similar to Drosophila M. Ab3-B gene
encoded by GenBank Accession Number L07835. The I.M.A.G.E.
Consortium clone ID number is 23931."
<..657
/note="similar to S-adenosylhomocysteine hydrolase"
/codon_start=1
/protein_id="AAB88189.1"
/db_xref="GI:2661083"
/translation="PRAMCGYGEVNGKCCALKALGAIVITEIDPICALQACMDGFR
VYKLEIVQDVIVITCTGNKRVYLAEGRLINLSCPTPTVLSITATTOALILEYNA
TGERVSOVDHYIMPDKGRVYLAEGRLINLSCPTPTVLSITATTOALILEYNA
PGRYKQDYLLPKKDEYVASHLPSPFAHLELTDDAKILGLNKNGPRPNRYR"

BASE COUNT 370 a 314 c 286 g 399 t
ORIGIN
Query Match 34.9% Score 644 DB 9: Length 1369;
Best Local Similarity 100.0% Pred. NO. 3.9e-119;
Matches 644 Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1201 TGTGTGCTATGTGAGTAGGCAAGGGCTGTGCTCTCAAGCTCTTGAGACAA 1260
DB 11 TGTGTGCTATGTGAGTAGGCAAGGGCTGTGCTCTCAAGCTCTTGAGACAA 70
QY 1261 TTGTCTACATACCGAATGACCCCATCTGTGCTCTGACAGGCTGATGAGAGGTTCA 1320
DB 71 TTGTCTACATACCGAATGACCCCATCTGTGCTCTGACAGGCTGATGAGAGGTTCA 130
QY 1321 GGGTGGTAAAGCTAAATGAAGTATCGGCAAGCGATGCTGAATTAATGACAGAGAA 1380

```

```

DB 131 GGGTGGTAAAGCTAAATGAAGTATCGGCAAGCTGCTGAATTAATGACAGAGAA 190
QY 1381 ATAGAATGATGATGACACGGGACCTTGATGATGAGAAACAGTTGATGATGCA 1440
DB 191 ATAGAATGATGATGACACGGGACCTTGATGATGAGAAACAGTTGATGATGCA 250
QY 1441 ATATGGGCCACTCCACACAGAAATGATGTGACACGCTCCGACTCCGGAGTACGT 1500
DB 251 ATATGGGCCACTCCACACAGAAATGATGTGACACGCTCCGACTCCGGAGTACGT 310
QY 1501 GGGAGGAGATGATGCTCTCAGTGGACCATGTCAATCGGCGAGATGGCAAGATTGTC 1560
DB 311 GGGAGGAGATGATGCTCTCAGTGGACCATGTCAATCGGCGAGATGGCAAGATTGTC 370
QY 1561 TCCTGGCAGAGGGTGTCTCAATCAATTTAGCTGTCCACAGTTCCACCTTTGTTCTGT 1620
DB 371 TCCTGGCAGAGGGTGTCTCAATCAATTTAGCTGTCCACAGTTCCACCTTTGTTCTGT 430
QY 1621 CCATCAGCCACACACACAGCGCTTTGGCACTGATAGACTCTATATGACACCGAGGGCC 1680
DB 431 CCATCAGCCACACACACAGCGCTTTGGCACTGATAGACTCTATATGACACCGAGGGCC 490
QY 1681 GATACAAAGCAGATGTGTACTTGTCTCTTAAGAAATGATGATAGCTTGCAGCTTGC 1740
DB 491 GATACAAAGCAGATGTGTACTTGTCTCTTAAGAAATGATGATAGCTTGCAGCTTGC 550
QY 1741 ATCTGCATCATTTGATGATGCCACCTTACAGAGCTGACAGATGACCAAAATATCTGG 1800
DB 551 ATCTGCATCATTTGATGATGCCACCTTACAGAGCTGACAGATGACCAAAATATCTGG 610
QY 1801 GACTCAACAAAATGGGCCATTCAACCTTAATTTACATAC 1844
DB 611 GACTCAACAAAATGGGCCATTCAACCTTAATTTACATAC 654

RESULT 15
BC003631
LOCUS Homo sapiens, similar to S-adenosylhomocysteine hydrolase-like 1, 1323 bp. mRNA linear PRI 12-JUL-2001
DEFINITION BC003631
ACCESSION BC003631.1 GI:13177702
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (26-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
CONTACT: nih-mgc project URL: http://mgc.ncl.nih.gov
CONTACT: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Keltman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: LNL Plate: 3 Row: D Column: 6.
Location/Qualifiers
1..1323
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

CDS

/clone="IMAGE:3010755"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_id="NIH_MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
<1..611
/codon_start=3
/product="Similar to S-adenosylhomocysteine hydrolase-1like
1"
/protein_id="AAH03631.1"
/db_xref="GI:13177703"
/translation="HEBALGIYITTEIDPICALQACMDGFRVYKLEIVTRQDVVIT
CTGNKNVYTRHLDKRNKNSCIYCMGHSNTEIDVTSLSRTEPELTERVRSQDVHWPD
GKRVYLAEGRLNLSCSTVPTFVLSITATQALALIELYNAPGRRYKODVYLLPKKK
DEVYASHLPSFDLHLELMDOKVYGLNKNKGFPRNYRY"

BASE COUNT 365 a 303 c 271 g 384 t
ORIGIN

Query Match 32.5%; Score 599; DB 9; Length 1323;
Best Local Similarity 100.0%; Pred. No. 4.2e-110;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1246 AAGCTCTGGAGCAATCTGTACATTTACCAATGACACCCCATCTGTGCTCGCAGGCCT 1305
|||||
Db 10 AAGCTCTGGAGCAATCTGTACATTTACCAATGACACCCCATCTGTGCTCGCAGGCCT 69
QY 1306 GCATGATGGGTTCAGGCTGTAAGTAATGAAGTATCCGGCAAGTCGATCTGTA 1365
|||||
Db 70 GCATGATGGGTTCAGGCTGTAAGTAATGAAGTATCCGGCAAGTCGATCTGTA 129
QY 1366 TAACCTGCACAGAAATAGAATAGTAGACACGGAGCACTTGATGCATGAAAAACA 1425
|||||
Db 130 TAACCTGCACAGAAATAGAATAGTAGACACGGAGCACTTGATGCATGAAAAACA 189
QY 1426 GTTGTATGATGCAATTTGGGCCCTCAACACAGAAATGATGTGACACAGCTCCGCA 1485
|||||
Db 190 GTTGTATGATGCAATTTGGGCCCTCAACACAGAAATGATGTGACACAGCTCCGCA 249
QY 1486 CTCCGAGACCTGACGAGGAGGAGTAGTCTCAGGTGAGACATGTCATCTGGCCAGATG 1545
|||||
Db 250 CTCCGAGACCTGACGAGGAGGAGTAGTCTCAGGTGAGACATGTCATCTGGCCAGATG 309
QY 1546 GCAACAGAGTTCCTCTGCGACAGAGGCTCTACTCAATTTAGCTGCTCCACAGTTC 1605
|||||
Db 310 GCAACAGAGTTCCTCTGCGACAGAGGCTCTACTCAATTTAGCTGCTCCACAGTTC 369
QY 1606 CCACCTTTGTTCTGTCACACACACACACAGGCTTTGGCACTGATAGAACTCTATA 1665
|||||
Db 370 CCACCTTTGTTCTGTCACACACACACACAGGCTTTGGCACTGATAGAACTCTATA 429
QY 1666 ATGCACCCGAGGGGAGATACAAAGAGATGTGACTTCTTCTTAAGAAAAATGATGAAT 1725
|||||
Db 430 ATGCACCCGAGGGGAGATACAAAGAGATGTGACTTCTTCTTAAGAAAAATGATGAAT 489
QY 1726 ACGTTGCAGCTTGATCTGCATCATTTGATGCCCACCTTACAGAGCTGACAGATGACC 1785
|||||
Db 490 ACGTTGCAGCTTGATCTGCATCATTTGATGCCCACCTTACAGAGCTGACAGATGACC 549
QY 1786 AAGCAAAATATCTGGGACTCAACAAAAATGGGCCATTCAAACTAATATTATACAGATAC 1844
|||||
Db 550 AAGCAAAATATCTGGGACTCAACAAAAATGGGCCATTCAAACTAATATTATACAGATAC 608

Search completed: April 21, 2003, 02:39:39
Job time : 4773.08 secs

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	Medline	PUBMED
AK014539	3234 bp.	mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4631427C17:homolog to PUTATIVE ADENOSYLMOMOCYSTEINASE (EC 3.3.1.1) (S-ADENOSYL-L- HOMOCYSTEINE HYDROLASE) (ADOCYASE), full insert sequence.	AK014539	GI:12852459	HTC; CAP trapper.	Mus musculus (strain:C57BL/6J) 0 day neonate skin cDNA to mRNA, clone_11b:RIKEN full-length enriched mouse cDNA library clone:4631427C17.	Mus musculus	1	Carninci, P. and Hayashizaki, Y.	High-efficiency full-length cDNA cloning	Meth. Enzymol. 303, 19-44 (1999)	9872923	10349636	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	Genome Res. 10 (10), 1617-1630 (2000)	20495374	

OY	798	ACACAGAAAGAGAGCTGACAGACACAGCGGCTGAGGGCTGGAGAGTTGACAGTGTGCGCTGGAG	857
Db	836	ACCCCTCATGAGATGGCTGCTGCTCTCCGAGAAAGTGGGTTTCCCGTCTTTGCTGGAAA	895
OY	858	GGCAGATCAGAAAGATGATCTTGGTGGTGTATTGACCCGCTGTGTGAACATGGATGGGTG	917
Db	896	GGAAATCAGAAAGATGATCTTGGTGGTGCATTTGACAGATGCGTGAATGTGGAGGCGTGG	955
OY	918	CAGGCCAATGATCCTGGATGATGGGGGAGACTTAACCATGTGGGTTTATTAAGAAAT	977
Db	956	CAGCCAAACATGATACGTGATGATGAGAGATCTGACACTGTGATTTATTAAGAAAT	1015
OY	978	CCAAACGTGTTTAAGAGATCCGAGGCATTGTGGAAGAGAGGCTGACTGTTTCACAG	1037
Db	1016	CCCAACATGTTTAAAAAAATCAAGGCATATGTCGAGAGAAAGTGTACTGAGATCCATAG	1075
OY	1038	CTGTATCAGCTCTCCAAAGCTGGGAAGCTGTGTCCGGCATGAGCTCAATGATCTT	1097
Db	1076	CTGTACCAACTGTCCTCAAGCTGGGAAGCTGTGTCTCCGCATACAGCTCAATGACCA	1135
OY	1098	GTTACCAACAGAAAGTTGATTAACCTGTACGTGTGCGGAGATTCATTTGGATGGCGCTG	1157
Db	1136	GTCCTTAAGCAGAAATTTGTAACCTCTACTGTGTCCGGAATCTATTTCTTGACGACTT	1195
OY	1158	AAGAGCCACACAGATGTGATGTTGGTGGGAAACAAGTGGTGTGTGGCTATGATG	1217
Db	1196	AAAGGACAAACAGATATGATGTTTGGTGGAAACAGAGTGGTGTGCTGATGGAAG	1255
OY	1218	GTAAGCAAGGCGCTGCTGTGCTGCTCCAAAGCTTTGAGCAATGTCTACATTACCGAA	1277
Db	1256	GTTGGAAAGGGGTCTGCTGCTGCTGTGAAAGCCATGGCTCAATGTGTACTGCTGACTAG	1315
OY	1278	ATCACCCCATCTGTGCTCTGACAGGCGCTGCATGATGATGGGTTCCAGGATGGTAAAGCTAAT	1337
Db	1316	ATTCACCCCATCTGTGCTCTGACAGGCGCTGCATGATGATGGGTTCCAGCTAATGAAAGCTAAT	1375
OY	1338	GAAATCATCCGGGCAAGTCATGTCTGTATTAATTACTTGACAGGAAATTAAGATTAATGACA	1397
Db	1376	GAAATCATCCGACAGGTCGACATGTTATTACTTGACAGGGAACAAGATGTGGTAAAC	1435
OY	1398	CGGAGCACTTGATTCGATGCAAGAAACACAGTTGATGTGATGCAATATGGGCACTCCAC	1457
Db	1436	AGAGAGCACTTGAGACCTATGAAAGAAATAGCTGATGTTGTAACTATGGCCATCTCAAC	1495
OY	1458	ACAGAAATCGATGTGACACAGCTCCGCACTCCGGAGCTGACGTGGAGCGAGTACG- TTC	1516
Db	1496	ACCGAAATGTGTGGCCACCTCGCGACACCCAAACTACCTGGGAGAGATTAAGATTTC	1555
OY	1517	TCAGGTGGACCATGTCATCTGGGCGAGT- GGCAGAAAGAGTGTGTCCTCGGAGAG- GGGT	1574
Db	1556	CCAAATTTGATATGTGATGATGAGCTGTGTTGGGCAAGAGATAGTATTGCTGGCAGAGGGCC	1615
OY	1575	CGCTACTCATTTGAGCTGCTCCACAGTTGCCACCTTTGTTCTGTCCATCCACGCCACA	1634
Db	1616	CGTTAATGAAATCACTCAGGTGCTCCACAGTGCCTACATTTGTGTCTCCATCACTGCTAC	1675
OY	1635	ACACAGCGTTTGGGACATGATGAACATCTATTAATGACCCCGAAGGGGAGATACAGAGAT	1699
Db	1676	ACTAGAGGCTCTCGCTGTGATGAACATTTAAAGTCTCGAAGGGTGTGCTATAAGAAAT	1735
OY	1695	GTTACTGCTTCTTAAGAAATGAGATGATAGTTGCCAGCTGTGCATCTGCATCATTT	1755
Db	1736	GTTACTGCTTCTTAAGAAATGAGATGATAGTTGCCAGCTGTGCATCTGCATCATTT	1799
OY	1755	GATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGGACTCAACAAAAT	1814
Db	1796	GATGCCACCTGACAGAGCTTACAGATGAACAAGCAAGATATCTGGGACTCAACAAAGAT	1855
OY	1815	GGGCGATCAAAACCTAATTAATACAGTA	1843
Db	1856	GGACCTTCAAGCTTAATCTACTACAGGTA	1864

RESULT 2	AL527928	938 bp	mRNA	linear	EST 13-FEB-2001
LOCUS	AL527928	LTI.NFL003.NBC3	Homo sapiens	cDNA clone CS0DC027YC19.5	
DEFINITION	AL527928	LTI.NFL003.NBC3	Homo sapiens	cDNA clone CS0DC027YC19.5	
ACCESSION	AL527928				
VERSION	AL527928.1	GI:12791421			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 938)				
JOURNAL	L1,W.B., Gruber,C., Jessee,J. and Polayes,D.				
COMMENT	Full-length cDNA libraries and normalization				
	Unpublished (2001)				
	Contact: Genoscope				
	Genoscope - Centre National de Sequencage				
	BP 191 91006 EVRY cedex - France				
	Email: sege@genoscope.cns.fr, Web : www.genoscope.cns.fr.				

```

FEATURES
source
Location/Qualifiers
1..938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODCC027Yc19"
/clone_11b="LTL_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
241 a 208 c 275 g 212 t 2 others

```

Query Match	Similarity	50.2%	Score 926.2	DB 9	Length 938
Best Local	Similarity	99.7%	Pred. Ns.3.9e-224		
Matches	936	Conservative	2	Mismatches	0
				Indels	1
				Gaps	1
QY	289	TCGGGGAGGAGCTGAAAGCAGGCGCCAGAGAGATCGAGGACCGCGAGAGTACTCTTCATG	348		
Db	1	TCGGGGAGGAGGTGAAGCAGAGGCGCCAGAGATCGAGGACCGCGAGAGTACTCTTCATG	60		
QY	349	CCACCGCTCACCAAGGCGCCCGAAGCAATTCAGTTTGTCTATGACATGCAAGATTC	408		
Db	61	CCACCGCTCACCAAGGCGCCCGAAGCAATTCAGTTTGTCTATGACATGCAAGATTC	120		
QY	409	CCAAATTCGCCACCAAAATCGGCCGAAGATCTTGTCTGCTGATCAGATCTCCGCA	468		
Db	121	CCAAATTCGCCACCAAAATCGGCCGAAGATCTTGTCTGCTGATCAGATCTCCGCA	180		
QY	469	CTGACAGCTACAGTTCAGCTGCATCCTACACAGATAGCTCTGATGATGAGTTTCC	528		
Db	181	CTGACAGCTACAGTTCAGTTCGCTGCATCCTACACAGATAGCTCTGATGATGAGTTTCC	240		
QY	529	GAGAGAGGAGCAAAACCAATCCCAAGGGCAGCAGCAATTTCTGTGAAGAACATCAGC	588		
Db	241	GAGAGAGGAGCAAAACCAATCCCAAGGGCAGCAGCAATTTCTGTGAAGAACATCAGC	300		
QY	589	AGGACGAATTTGACCGCGGGAGATTGAATTGACAGAGCAAGACATGTCCTGATTT	648		
Db	301	AGGACGAATTTGAGCCGGGAGATTGAGATTGACAGAGCAAGACATGTCCTGATTT	360		
QY	649	CACCTAGGAAAGCTGCTCAGAGGGGAGAAAGCCCTTGGCTGCTTAATAATAGTGGCTGA	708		
Db	361	CACCTAGGAAAGCTGCTCAGAGGGGAGAAAGCCCTTGGCTGCTTAATAATAGTGGCTGA	420		

QY 709 CACATACACAGCCAGACGCGTGTGATGATGAGACACTGTGCGCTGGGGCTCAGT 768
 DB 421 CACATACACAGCCAGACGCGTGTGATGATGAGACACTGTGCGCTGGGGCTCAGT 480
 QY 769 GCGCGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 828
 DB 481 GCGCGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 QY 829 AGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888
 DB 541 AGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 889 TTACCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
 DB 601 TTACCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 949 ACTTACCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008
 DB 661 ACTTACCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 QY 1009 TGGAGAGAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
 DB 721 TGGAGAGAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 QY 1069 GTGTCCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1128
 DB 781 GTGTCCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 QY 1129 GCTGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188
 DB 841 GCTGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 899
 QY 1189 AACAGT 1227
 DB 900 AACAGT 938
 RESULT 3
 LOCUS AL551097 921 bp mRNA linear EST 16-FEB-2001
 DEFINITION AL551097 L1_NFL006_P12 Homo sapiens cDNA clone CS0D1066Y07 5
 prime, mRNA sequence.
 ACCESSION AL551097
 VERSION AL551097.1 GI:1288715
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 921)
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 source
 1. 921
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1066Y07"
 /location="L1_NFL006_P12"
 /issue_type="placenta"
 /note="Vector: pCMVSPORT 6; site_1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with NotI and
 cloned into the NotI and EcoRV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 life technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@life.techn.com URL :
 http://fulllength.invitrogen.com
 BASE COUNT 217 a 231 c 302 g 170 t 1 others
 ORIGIN
 Query Match 48.7%; Score 897.4; DB 9; Length 921;
 Best Local Similarity 99.6%; Pred. No. 8e-217;
 Matches 909; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 QY 69 GCGTGGGCTGCGGACACAGACAGCGCTGGGCGACAGCTCAGAAAGCCGACAGCTC 128
 DB 1 GCGTGGGCTGCGGACACAGACAGCGCTGGGCGACAGCTCAGAAAGCCGACAGCTC 60
 QY 129 GAGCGAGGCGCGGACAGAGGCTGGGATGCTGTCGAGAGGCGCGCGCGCGGCGAGCC 188
 DB 61 GAGCGAGGCGCGGACAGAGGCTGGGATGCTGTCGAGAGGCGCGCGCGCGGCGAGCC 120
 QY 189 GGGCGGCGCGGACAGAGGCGGGAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 248
 DB 121 GGGCGGCGCGGACAGAGGCGGGAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 QY 249 GGGGGAATGTGATGCTGTGACGCGATCCGCTGCCGGGGTGGGGAGAGCTGAAGCG 308
 DB 181 GGGGGAATGTGATGCTGTGACGCGATCCGCTGCCGGGGTGGGGAGAGCTGAAGCG 240
 QY 309 GCCAAGAGATCCAGAGAGCGCGGAGATCTCTCATATGGCCACCTGCACCAAGCGCGCC 368
 DB 241 GCCAAGAGATCCAGAGAGCGCGGAGATCTCTCATATGGCCACCTGCACCAAGCGCGCC 300
 QY 369 AAGAGAAATCCAGATTGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 428
 DB 301 AAGAGAAATCCAGATTGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 429 GCGCGAAGATCTTGTCTGCTGCTGATCTCAGATCTCAGATCTCAGATCTCAGATCTCAG 488
 DB 361 GCGCGAAGATCTTGTCTGCTGCTGATCTCAGATCTCAGATCTCAGATCTCAGATCTCAG 420
 QY 489 GCATCTTACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 548
 DB 421 GCATCTTACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 549 TCCAGAGGCGAGCAGCAATTTCTGTGTGAAGAAATCAAGAGGAGAGATTGGAGCGCGG 608
 DB 481 TCCAGAGGCGAGCAGCAATTTCTGTGTGAAGAAATCAAGAGGAGAGATTGGAGCGCGG 540
 QY 609 GAGATTGAGATTGAGAGAGAGAGATGCTGCTGATTTCTCAGAGAAACGTGCTCAG 668
 DB 541 GAGATTGAGATTGAGAGAGAGAGATGCTGCTGATTTCTCAGAGAAACGTGCTCAG 600
 QY 669 GGGGAGAGCGCTTGGCTGGTGGCTTAAATAGTGGGCTGTACACATCAGAGCCAGACA 728
 DB 601 GGGGAGAGCGCTTGGCTGGTGGCTTAAATAGTGGGCTGTACACATCAGAGCCAGACA 660
 QY 729 GCGGTGTGATTGAGACACTGTGCGCTGGGGGCTCAGTCCGCTGGTGTGTTGATC 788
 DB 661 GCGGTGTGATTGAGACACTGTGCGCTGGGGGCTCAGTCCGCTGGTGTGTTGATC 720
 QY 789 ATCTACTCACTCAGAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 848
 DB 721 ATCTACTCACTCAGAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 780
 QY 849 GCTTGAAGAGGCGAGTGAAGAGATGATCTGAGGGGTGATGAGCCCTGTGTAAATG 908
 DB 781 GCTTGAAGAGGCGAGTGAAGAGATGATCTGAGGGGTGATGAGCCCTGTGTAAATG 840
 QY 909 GATGGTGGCAGGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 968
 DB 841 GATGGTGGCAGGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 899
 QY 969 AAGAGATATCCAA 981
 DB 900 AAGAGATATCCAA 912

RESULT 4	AL520704	LOCUS	AL520704	DEFINITION	AL520704 LTR_NFL004_NBC2 Homo sapiens CDNA clone CS0DB002YL22 5	891 bp	mRNA	linear	EST 13-FEB-2001
ACCESSION	AL520704	VERSION	AL520704.1	GI:12784197					
SOURCE	ORGANISM	human.							
REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT					
1	(bases 1 to 891)	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.							
1	L.N.B., Gruber,C., Jessee,J. and Polayes,D.	Full-length cDNA libraries and normalization	unpublished (2001)						
Contact:	Genoscope								
Genoscope	- Centre National de Sequencage								
BP 191	91006 EVRY cedex - France								
Email:	segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.								
Location/Qualifiers									
1.	891								
/organism="Homo sapiens"									
/db_xref="taxon:9606"									
/clone="CS0DB002YL22"									
/clone_11b="LTR_NFL004_NBC2"									
/sex="male"									
/tissue.type="neuroblastoma cells"									
/lab_host="DH10B"									
/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA									
was primed with a NotI-OLIGO(dt) primer. Five prime end									
enriched, double-stranded cDNA was digested with Not I and									
cloned into the Not I and Eco RV sites of the pCMVSPORT 6									
vector. Library was normalized. Library was constructed									
by Life Technologies. Contact: Feng Liang Life									
Technologies, a division of Invitrogen 9800 Medical Centre									
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610									
8371 Email : filiang@lifetech.com URL :									
http://fulllength.invitrogen.com"									
BASE COUNT	194 a	238 c	299 g	158 t	2 others				
ORIGIN									
Query Match	47.3%	Score 873;	DB 9;	Length 891;					
Best Local Similarity	99.5%;	Pred. No. 1.2e-210;							
Matches 873; Conservative	2;	Mismatches 2;	Indels	0;	Gaps	0;			
1	GGCGGGGCGAGTCCGAGCTCGGAGCTGCTGCTTCTGCTCTTGTGGCCACCTCGTGGT	60							
15	GGCGGGGCGAGTCCGAGCTCGGAGCTGCTGCTTCTGCTCTTGTGGCCACCTCGTGGT	74							
61	GGCGGGGCGAGTCCGAGCTCGGAGCTGCTGCTTCTGCTCTTGTGGCCACCTCGTGGT	120							
75	GGCGGGGCGAGTCCGAGCTCGGAGCTGCTGCTTCTGCTCTTGTGGCCACCTCGTGGT	134							
121	GGCGGGGCGAGTCCGAGCTCGGAGCTGCTGCTTCTGCTCTTGTGGCCACCTCGTGGT	180							
135	GGCGGGGCGAGTCCGAGCTCGGAGCTGCTGCTTCTGCTCTTGTGGCCACCTCGTGGT	194							
181	GGCGGGGCGAGTCCGAGCTCGGAGCTGCTGCTTCTGCTCTTGTGGCCACCTCGTGGT	240							
195	GGCGGGGCGAGTCCGAGCTCGGAGCTGCTGCTTCTGCTCTTGTGGCCACCTCGTGGT	254							
241	GGCGGGGCGAGTCCGAGCTCGGAGCTGCTGCTTCTGCTCTTGTGGCCACCTCGTGGT	300							
255	GGCGGGGCGAGTCCGAGCTCGGAGCTGCTGCTTCTGCTCTTGTGGCCACCTCGTGGT	314							
301	GGCGGGGCGAGTCCGAGCTCGGAGCTGCTGCTTCTGCTCTTGTGGCCACCTCGTGGT	350							
315	GGCGGGGCGAGTCCGAGCTCGGAGCTGCTGCTTCTGCTCTTGTGGCCACCTCGTGGT	374							
361	GGCGGGGCGAGTCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	420							

	Db	375	AGGGGCCCCAAGAAGAAATTCAGATTGTGATGCATCGACGAGTCCACCATAATTCGCCA	434
OY	421	CCAAAATCTGGCCGAAGATCTTTGTCTCGCTGATCTCACAATCTCTCCACTACAGCTACA	480	
Db	435	CCAAAACTGGCCGAAGATCTTTGTCTCGCTGATCTCACAATCTCTCCACTACAGCTACA	494	
OY	481	GTTACGCTGCAATCTCACACATATAGCTGTGATGATGAGTTTCTCCCGAGAGAACAGC	540	
Db	495	GTTACGCTGCAATCTCACACATATAGCTGTGATGATGAGTTTCTCCCGAGAGAACAGC	554	
OY	541	AAACCACTCCAAGGGCAGCAGCATTTCTGTGTGAAGACATCAAGCAGCAGCAATTTG	600	
Db	555	AAACCACTCCAAGGGCAGCAGCATTTCTGTGTGAAGACATCAAGCAGCAGCAATTTG	614	
OY	601	GACGCCGGAGATTGAGATTGACAGCAGACATGCTGCTCTGATTTTCACTCAGGAAC	660	
Db	615	GACGCCGGAGATTGAGATTGACAGCAGACATGCTGCTCTGATTTTCACTCAGGAAC	674	
OY	661	GTGCTCAGGGGAGAGAGCCCTTGCTGTGTCTTAATAATGTGAGGCTGTACACATCAAG	720	
Db	675	GTGCTCAGGGGAGAGAGCCCTTGCTGTGTCTTAATAATGTGAGGCTGTACACATCAAG	734	
OY	721	CCCACAGCAGCGGTGTTGATTGAGACACTGTGCTGCGGCGCTCAGTAGCCGCTGTTG	780	
Db	735	CCCACAGCAGCGGTGTTGATTGAGACACTGTGCTGCGGCGCTCAGTAGCCGCTGTTG	794	
OY	761	CTTGTAACATCTACTCAACTCAGAAATGAATGCTGACGACACTGGCTAGGCTGAGTTG	840	
Db	795	CTTGTAACATCTACTCAACTCAGAAATGAATGCTGACGACACTGGCTAGGCTGAGTTG	854	
OY	841	CAGTGTGCTTGGAAAGGCGAGTCAGAAAGATGACTT	877	
Db	855	CAGTGTGCTTGGAAAGGCGAGTCAGAAAGATGACTT	891	
RESULT 5				
LOCUS	BMA49470	1042 bp	mRNA	linear EST 05-FEB-2002
DEFINITION	AGENCOURT_6400881 NIH_MGC_67 Homo sapiens CDNA clone IMAGE:5495677			
ACCESSION	BMA49470			
VERSION	BMA49470.1 GI:18498510			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	NIH-MGC http://mgc.nci.nih.gov/.			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgrabs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LNL12117 row: m column: 06 High quality sequence stop: 694. Location/Qualifiers 1..1042 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5495677" /clone_lid="NIH_MGC_67" /tissue_type="retinoblastoma" /lab_host="DH10B (phage-resistant)" /note="Organ: eye; Vector: pCMV-Sport6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT Average insert size 1.75 kb. Library constructed by Life			
FEATURES	Source			

BASE COUNT 267 a 219 c 309 g 240 t 7 others
 ORIGIN

Query Match 46.8%; Score 862.6; DB 13; Length 1042;
 Best Local Similarity 96.9%; Pred. No. 5.9e-208;
 Matches 908; Conservative 0; Mismatches 25; Indels 4; Gaps 3;

532 AGAAGCAGCAACCACTCCAGGGGAGAGCAATTTCTGTGTAACATCAAGCAG 591
 1 AGAAGCAGCAACCACTCCAGGGGAGAGCAATTTCTGTGTAACATCAAGCAG 60
 592 CAGATTGTGAGCGCGGAGATTGAGATTGCAGCAAGCAATGCTGCTGATTTTCA 651
 61 CAGATTGTGAGCGCGGAGATTGAGATTGCAGCAAGCAATGCTGCTGATTTTCA 120
 652 TCAGGAACGTGCTCAGGGGAGAGCCCTTGCTGCTGCTAAATAGTGGCTGTAC 711
 121 TCAGGAACGTGCTCAGGGGAGAGCCCTTGCTGCTGCTAAATAGTGGCTGTAC 180
 712 ACATCAGCCGAGACAGCGGTGTTGATTGAGACACTCTGCTGCGGGGCTCAGT 771
 181 ACATCAGCCGAGACAGCGGTGTTGATTGAGACACTCTGCTGCGGGGCTCAGT 240
 772 GCTGCTGCTGTTGATCACTCACTCAAGTAAGTGAAGTGAAGTGAAGTGAAG 831
 241 GCTGCTGCTGTTGATCACTCACTCAAGTAAGTGAAGTGAAGTGAAGTGAAG 300
 832 CTGAGATTGAGTGTGCTGCTGAGAGGCGAGTCAGAAAGATGACTTGTGCTGAT 891
 301 CTGAGATTGAGTGTGCTGCTGAGAGGCGAGTCAGAAAGATGACTTGTGCTGAT 360
 892 ACCGCTGTGGAACATGATGAGGCTGAGGCGCAACATGATCTGATGATGAGG 951
 361 ACCGCTGTGGAACATGATGAGGCTGAGGCGCAACATGATCTGATGATGAGG 420
 952 TAACCCACTGGTTTATTAAGAAATATCCAAAGCTTTAAGAAATCCGAGCA 1011
 421 TAACCCACTGGTTTATTAAGAAATATCCAAAGCTTTAAGAAATCCGAGCA 480
 1012 AAGAGCGTGAATGCTGCTCAGAGCTGATCAGCTCTCCAAAGCTGGAG 1071
 481 AAGAGCGTGAATGCTGCTCAGAGCTGATCAGCTCTCCAAAGCTGGAG 540
 1072 TTCCGGCATGAACGTAAAGATTTCTTTACCAAGAAAGTTGATTAATCTG 1131
 541 TTCCGGCATGAACGTAAAGATTTCTTTACCAAGAAAGTTGATTAATCTG 600
 1132 GCCGAGATCCATTTGATGAGCTGAGAGGAGCAGACAGATGATGTTGGTGA 1191
 601 GCCGAGATCCATTTGATGAGCTGAGAGGAGCAGACAGATGATGTTGGTGA 660
 1192 AAGTGTGTGTGTGCTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1251
 661 AAGTGTGTGTGTGCTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 720
 1252 TTGGAGCAATTTGCTACATTAACGAAATCGAACCCATCTGTGCTCTCAG 1311
 721 TTGGAGCAATTTGCTACATTAACGAAATCGAACCCATCTGTGCTCTCAG 780
 1312 ATGGG--TTAGGGTGTAAAGCTAAATGAAGTCAATCCGAGTCAATGCTA 1369
 781 ATGGGTTTCAAGGTTGTTAAAGCTAAATGAAGTCAATGCTGCTGCTA 840
 1370 TTGCACAGCAATTAAGATAGTGAACAGGAGCACTT--GGATGCGATGAAA 1428
 841 TTGCACAGCAATTAAGATAGTGAACAGGAGCACTTGGGATGGATGAAA 900
 1429 GTATC-GTATGCAATATGGGCACTCCACACAGAAA 1464
 901 GTATGCTATGCAATATGGGCACTCCACACAGAAA 937

RESULT 6
 AL519571 865 bp mRNA linear EST 13-FEB-2001
 LOCUS
 DEFINITION AL519571.L1.NFL004.NBC2 Homo sapiens cDNA clone CS0DB004Y104.5
 ACCESSION AL519571
 VERSION AL519571.1 GI:12783064
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 865)
 REFERENCE
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polyses, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 source
 1. 865
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="L1.NFL004.NBC2"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 193 a 231 c 284 g 155 t 2 others
 ORIGIN

Query Match 46.4%; Score 856.2; DB 9; Length 865;
 Best Local Similarity 99.2%; Pred. No. 2.2e-206;
 Matches 856; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

8 GCAAGTGGAGCTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 67
 1 GCAAGTGGAGCTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
 68 TGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 127
 61 TGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 128 GCAAGTGGAGCTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 187
 121 GCAAGTGGAGCTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 188 CGGCGGGCGCCAGAGGGGAGAGAGCGGGGAGCGGCTGCTGCTGCTGCTGCTG 247
 181 CGGCGGGCGCCAGAGGGGAGAGAGCGGGGAGCGGCTGCTGCTGCTGCTGCTG 240
 248 CGGCGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
 241 CGGCGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 308 GCGCAAGAGATGAGAGCGCGAGAGTACTCTTATGAGCCACGCTCACCAAGG 367
 301 GCGCAAGAGATGAGAGCGCGAGAGTACTCTTATGAGCCACGCTCACCAAGG 360
 368 CAGAGCAAAATCAGTTTGTGATGACATGACAGAGTTCACCAAAATCCCAAC 427
 361 CAGAGCAAAATCAGTTTGTGATGACATGACAGAGTTCACCAAAATCCCAAC 420

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TMAGE:5578170"
/clone_1lb="NH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site_1: Notti,
Site_2: Sall; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
technologies."

```

[illegible]

RESULT	10
AL525407	
LOCUS	
DEFINITION	839 bp mRNA linear EST 13-FEB-2001 AL525407.LTT.NFL003.NBC3 Homo sapiens CDNA clone CSDDC01LYD21.5
ACCESSION	prime, mRNA sequence.
VERSION	AL525407
KEYWORDS	AL525407.1 GI:12788900
SOURCE	EST.
ORGANISM	human. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE	1 (bases 1 to 839)
AUTHORS	Ll.W.B., Gruber,C., Jeessee,J. and Polayzes,D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES	
source	Location/Qualifiers 1..839 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="CS0DC0117D21" /clone_id="LTL_NFL003_NBC3" /sex="male" /tissue_type="neuroblastoma cells" /lab_host="BDH108" /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact : Feng liang life technologies, a division of Invitrogen 9600 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT	183 a 226 c 270 g 151 t 9 others
ORIGIN	
Query Match	44.1%; Score 814; DB 9; Length 839;
Best Local Similarity	98.1%; Pred. No. 1,2e-195;
Matches 824; Conservative 9; Mismatches 6; Indels 1; Gaps 1;	
QY	9 CAGGTGGAGAGCTGGAGAGCTGCTGCTTGTGCTTGTGTGGACACCGTCGCTGTGGCT 68
Db	1 CAGGTGGAGAGCTGGAGAGCTGCTGCTTGTGCTTGTGTGGACACCGTCGCTGTGGCT 60
QY	69 GCCTTGGGCTCCGGAACAGACAAGGCGTGGGCGACAGACCTCAGAAAGCCGACGACTC 128
Db	61 GCCTTGGGCTCCGGAACAGACAAGGCGTGGGCGACAGACCTCAGAAAGCCGACGACTC 120
QY	129 GACGCGAGGGCGCGGAGAGAGGTGGCGATGCGCTGTGGAAGGCGCGCGCGGAGAGC 188
Db	121 GACGCGAGGGCGCGGAGAGAGGTGGCGATGCGCTGTGGAAGGCGCGCGCGGAGAGC 180
QY	189 GGGCGGGGCGCCAGAGAGGGGAAAAGAGCGGGGCGGCGGCTACGCGCTGCGCGGGCGGC 248
Db	181 GGGCGGGGCGCCAGAGAGGGGAAAAGAGCGGGGCGGCGGCTACGCGCTGCGCGGGCGGC 240
QY	249 GGGGGAATGTGAGAGCCCTGAGCGCGATGCGCGCGCGGGGAGTGGAGAGAGAGTGAAGAG 308
Db	241 GGGGGAATGTGAGAGCCCTGAGAGCGATGCGCGCGCGGGGAGTGGAGAGAGAGTGAAGAG 300
QY	309 GCCAAGGAGATCGAGAGAGCGCGAAGATATCTTCATGAGGCGACCGTCAACCAAGGCGCC 368
Db	301 GCCAAGGAGATCGAGAGAGCGCGAAGATATCTTCATGAGGCGACCGTCAACCAAGGCGCC 360
QY	369 AAGAAGCAATTCAGATTGCTGATGATCATGCAAGAGATTACCAAAATTCGCCACCAAACT 428
Db	361 AAGAAGCAATTCAGATTGCTGATGATCATGCAAGAGATTACCAAAATTCGCCACCAAACT 420
QY	429 GGGCGAAGATTTGTGCGGTGCTCATGCAAGATTCGCAATGACATGACAGTAAAGTTCAAGT 488
Db	421 GGGCGAAGATTTGTGCGGTGCTCATGCAAGATTCGCAATGACATGACAGTAAAGTTCAAGT 480
QY	489 GCATCTTACACAGATAGCTGTGATGATAGGTTTCCCGCAGAGAAACCAACCAAC 548
Db	481 GCATCTTACACAGATAGCTGTGATGATAGGTTTCCCGCAGAGAAACCAACCAAC 540
QY	549 TCCAAGGGCGACGACAAATTTCTGTGTGAAGACATCAAGACGAGAAATTTGAGCGCGG 608
Db	541 TCCAAGGGCGACGACAAATTTCTGTGTGAAGACATCAAGACGAGAAATTTGAGCGCGG 600

QY	609	GAGATTGAGATTGGACAGACAGACAAATGCTGTGCTGATTCTACTAGAGAAAGTGTCTAG	668
Db	601	GGATTTGAGATTGGACAGACAGACAAATGCTGTGCTGATTCTACTAGAGAAAGTGTCTAG	660
QY	669	GGGAGAGAGCCCTTGGCTGCTGCTAAATATAGAGGCTGTACACATACAGACCCAGACA	728
Db	661	GGGAGAGAG-CTTGGCTGCTGCTAAATATAGAGGCTGTACACATACAGACCCAGACA	719
QY	729	GGGGTGTGATTGAACACTCTGTGCTCCCTGGGGGCTCAGTCCGCTGTCTGTGAAC	788
Db	720	GGGGTGTGATTGAACACTCTGTGCTCCCTGGGGGCTCAGTCCGCTGTCTGTGAAC	779
QY	789	ATCTACTCACTCAGAAATGAGTACTGTGAGACAGCGCGTGGAGCTGGAGTTGCACTGC	848
Db	780	ATCTACTCACTCAGAAATGAGTACTGTGAGACAGCGCTGTGAGCTGGAGTTGCACTGTTCC	839
RESULT	11		
LOCUS	BM453920	867 bp	mRNA linear EST 05-PEB-2002
DEFINITION	AGENCOURT_6402958 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:552895		
ACCESSION	BM453920		
VERSION	BM453920.1	GI:18502949	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 867)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabs-remail.nih.gov		
	Tissue Procurement: ATCC/DCTP/DRP		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	plate: LHAM12206 row: 1 column: 20		
	High quality sequence stop: 670.		
FEATURES			
source	Location/Qualifiers		
	1..867		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5528955"		
	/clone_lid="NIH_MGC_72"		
	/tissue_type="melanotic melanoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: skin; Vector: pCMV-SPORT6; site_1: NotI; site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."		
BASE COUNT	222 a 201 c 241 g 202 t	1 others	
ORIGIN			
Query Match:	43.88; Score 807.4; DB 13; Length 867;		
Best Local Similarity	97.78; Pred. No. 5.6e-194;		
Matches 840; Conservative	0; Mismatches 17; Indels 3; Gaps 2;		
QY	338	CCGGAAGTACTCTTATATGGCCACCGTACACAGGGGCCCAAGAGCAATCCAGTTTG	387
Db	1	CCGGAAGTACTCTTATATGGCCACCGTACACAGGGGCCCAAGAGCAATCCAGTTTG	60
QY	388	CTGATGACATCGAGAGTTCACCAAAATCCCCACCAAAAGCGGCGAAGATCTTGTCTC	447
Db	61	CTGATGACATCGAGAGTTCACCAAAATCCCCACCAAAAGCGGCGAAGATCTTGTCTC	120
QY	448	GCTGATCTCAGCTCCGCCACTGACAGCTACAGTTCAGTGCATCTTACACAGATAGCT	507

[illegible]

QY 1508 AGTACGTTCTCAGGTGGACCATGTCTATCTGGCCAGATGGCAACGAGTTGTCTCTCTGGC 1509

9.1
9.2

Db 61 r

Db 61 r


```

Db 602 TTACCGCTGTGTGAACATGATGGGTGGAGGCAACATGATCTGTGATGGGGAG 661
QY 949 ACTT-AACCATCTGGTGTATTAAGAAATATCCAAAGCTGTTTAAGAAGAT-CCGAGGCAT 1006
Db 662 ACTTAACCTACTGGTGTATTAAGAAATATCCAAAGCTGTTTAAGAAGATCCGAGGCAT 721
QY 1007 TGTGAAGAGAGGCTGATCTGTGTACAGGCTGTATCACTCTCCAAAGCTGGGAAGCT 1066
Db 722 TGTGAAGAGAGGCTGATCTGTGTACAGGCTGTATCACTCTCCGCAAGCTGGGAAGCT 781
QY 1067 CTGTGTCCCGCCATGATCAATGATCTGTGTACCAAGAGATTTGATTAAGCTGTA 1126
Db 782 CTGTGTCCCGCCATGATCAATGATCTGTGTACCAAGAGATTTGATTAAGCTGTA 841
QY 1127 CTGTGTCCCGCCATGATCAATGATCTGTGTACCAAGAGATTTGATTAAGCTGTA 1141
Db 842 ATGCTGCCGGGAATC 856

```

```

RESULT 15
LOCUS AU126038 764 bp mRNA linear EST 01-AUG-2002
DEFINITION AU126038 NT2RM4 Homo sapiens cDNA clone NT2RM4002613 5', mRNA
sequence.
ACCESSION AU126038
VERSION AU126038.1 GI:10950754
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE 1 (bases 1 to 764)
AUTHORS Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
Isogai,T.
HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
Y., Sugano,S., Isogai,T.)
Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix
Research Institute: cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers

```

```

FEATURES
source
1. 764
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM4002613"
/clone_1b="NT2RM4"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/note="Vector: PME18FL3; mRNA from uninduced NT2 neuronal
precursor cells"

```

```

BASE COUNT 193 a 157 c 228 g 183 t 3 others
ORIGIN

```

```

Query Match 39.7%; Score 732.8; DB 9; Length 764;
Best Local Similarity 99.1%; Pred. No. 4.6e-175;
Matches 756; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

```

```

QY 472 ACAGCTACAGTTCAGCTGCATCTTACACAGATAGCTGTGATGAGGTTTCTCCCGAG 531
Db 1 ACAGCTACAGTTCAGCTGCATCTTACACAGATAGCTGTGATGAGGTTTCTCCCGAG 60
QY 532 AGAAGAGCAAAACCACTCCAGGAGGAGCAATTTCTGTGAGAAACATCAAGCAG 591

```

```

Db 61 AGAAGAGCAAAACCACTCCAGGAGGAGCAAGCAATTTCTGTGAGAAACATCAAGCAG 120
QY 592 CAGAATTTGAGAGCGCGGAGATTTGATTTGATTCAGAGCAAGACATGCTGCTGTATTTAC 651
Db 121 CAGAATTTGAGAGCGCGGAGATTTGATTTGATTCAGAGCAAGACATGCTGCTGTATTTAC 180
QY 652 TCAGAAACGTCCTCAGAGGAGAGAGAGCCCTTGGCTGCTAAATAGTGGCTGTACAC 711
Db 181 TCAGAAACGTCCTCAGAGGAGAGAGAGCCCTTGGCTGCTAAATAGTGGCTGTACAC 240
QY 712 ACATCACAGCCAGACAGCGGCTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 771
Db 241 ACATCACAGCCAGACAGCGGCTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 300
QY 772 GCTGCTGCTGTTTGAACATCTACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 831
Db 301 GCTGCTGCTGTTTGAACATCTACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 360
QY 832 CTGAGTTGCAAGTGTGCTTGGTGAAGGCGAGTCAGAAAGTACTTGTGCTGTATTTG 891
Db 361 CTGAGTTGCAAGTGTGCTTGGTGAAGGCGAGTCAGAAAGTACTTGTGCTGTATTTG 420
QY 892 ACCGCTGTGAACATGATGGGTGGGAGGCGCAACATGATGATGATGATGATGATGATGATG 951
Db 421 ACCGCTGTGAACATGATGGGTGGGAGGCGCAACATGATGATGATGATGATGATGATGATG 480
QY 952 TAACCCACTGGGTTTATTAAGAAATATCCAAAGCTGTTTAAGAAATATCCAGAGCTGTGG 1011
Db 481 TAACCCACTGGGTTTATTAAGAAATATCCAAAGCTGTTTAAGAAATATCCAGAGCTGTGG 540
QY 1012 AAGAGAGCTGACTGCTGTTCACAGGCTGTATCACTCTCCAAAGCTGGAGAGCTGTG 1071
Db 541 AAGAGAGCTGACTGCTGTTCACAGGCTGTATCACTCTCCAAAGCTGGAGAGCTGTG 600
QY 1072 TTCGGGCGATGACGTCATGATGATCTGTTCACAAAGAGAGTTGATGATGATGATGATG 1131
Db 601 TTCGGGCGATGACGTCATGATGATCTGTTCACAAAGAGAGTTGATGATGATGATGATG 659
QY 1132 GCCGAGATTCATTTTGATGGCCCTGAGAGGAGCCACAGATGTGATGTTGTGGGAAC 1191
Db 660 GCCGAGATTCATTTTGATGGCCCTGAGAGGAGCCACAGATGTGATGTTGTGGGAAC 719
QY 1192 AAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1233
Db 720 AAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 762

```

Search completed: April 21, 2003, 03:59:15
 Job time: 2497.43 secs

PN	W09814562-11.	
XX		
PD	09-APR-1998.	
XX		
PE	06-OCT-1997;	97W0-NZ00133
XX		
PR	04-OCT-1996;	96NZ-0299507
XX		
PA	(HART/) HART D N J.	
XX		
PI	Hart DNJ;	

XX WPI: 1998-240074/21.
DR P-PSDB; AARS6097.
XX

PT New isolated S-adenosyl-L-homocysteine hydrolase enzyme - is used to
PT develop products which can be used in the treatment of e.g.
PT auto-immune disease, transplantations or cancers
XX

PS Claim 7; Fig 1; 33pp; English.

XX This is the nucleotide sequence of the DP4b9.3 enzyme which has
CC S-adenosyl-L-homocysteine hydrolase (AHCY)-type activity. Its
CC products can be used to identify substances which have a stimulatory
CC or inhibitory effect on the enzyme activity. Such substances can be
CC used to modulate dendritic cell (DC) function and for immunomodulation.
CC They can be used in the treatment of e.g. autoimmune diseases,
CC transplantations or cancers. The products can also be used for
CC detection and diagnosis.

XX Sequence 2563 BP; 646 A; 604 C; 677 G; 636 T; 0 other;

Query Match 100.0%; Score 1844; DB 19; Length 2563;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1844; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCGGGGAGGTCGAGCTCGGAGCTGCTTGTGTTCTGTTGGCCACCGTGGT 60
DB 1 GGGCGGGGAGGTCGAGCTCGGAGCTGCTTGTGTTCTGTTGGCCACCGTGGT 60
QY 61 GTCCGGCTCTTGGGCTCCGAAACAGACAGAGGCTGGGCAACGACCTCAGAACCCGA 120
DB 61 GTCCGGCTCTTGGGCTCCGAAACAGACAGAGGCTGGGCAACGACCTCAGAACCCGA 120
QY 121 CGCAGCTCAGCAGAGGCGGCGAGAGAGGTGGCGATGCGGTGTGAGAGGCGCGCGC 180
DB 121 CGCAGCTCAGCAGAGGCGGCGAGAGAGGTGGCGATGCGGTGTGAGAGGCGCGCGC 180
QY 181 GGGCAGGCGGGGGGCGGCGGAGAGAGGAGGAGGCGGGGCGGCTAGCGCGGCGC 240
DB 181 GGGCAGGCGGGGGGCGGCGGAGAGAGGAGGAGGCGGGGCGGCTAGCGCGGCGC 240
QY 241 GGGCGGGCGGGGGAATGTGATGCTGACGCGATGCGCGTCCCGGGGTGCGGGAGGAGC 300
DB 241 GGGCGGGCGGGGGAATGTGATGCTGACGCGATGCGCGTCCCGGGGTGCGGGAGGAGC 300
QY 301 TGAAGCAGGCGCAGAGATGAGAGAGCGGAGAGAGTACTCTTCATGAGCGCACCGTACCA 360
DB 301 TGAAGCAGGCGCAGAGATGAGAGAGCGGAGAGAGTACTCTTCATGAGCGCACCGTACCA 360
QY 361 AGGCGCGCAGAGAAATCCAGTTTGTGATGACATGCAAGAGTTCACCAAAATCCCA 420
DB 361 AGGCGCGCAGAGAAATCCAGTTTGTGATGACATGCAAGAGTTCACCAAAATCCCA 420
QY 421 CCAAAAGTGGCGAGAGATTTGTCTGCTGATCTCAAGTCCATGACAGTGA 480
DB 421 CCAAAAGTGGCGAGAGATTTGTCTGCTGATCTCAAGTCCATGACAGTGA 480
QY 481 GTTCAAGTCTCTACACAGATAGTGTGATGAGAGTTTCCCGAGAGAGAGCAGC 540
DB 481 GTTCAAGTCTCTCTACACAGATAGTGTGATGAGAGTTTCCCGAGAGAGAGCAGC 540
QY 541 AAAGCAATCCAGAGGAGCAGCAATTTCTGTGTAAGAACATCAAGCAGAGAGATTGG 600
DB 541 AAAGCAATCCAGAGGAGCAGCAATTTCTGTGTAAGAACATCAAGCAGAGAGATTGG 600
QY 601 GAGCGCGGAGAGTGTGATGAGAGAGAGAGATGCTCTGATTTACTCAGAGAAAC 660
DB 601 GAGCGCGGAGAGTGTGATGAGAGAGAGAGATGCTCTGATTTACTCAGAGAAAC 660
QY 661 GTGCTCAGAGGAG 720
DB 661 GTGCTCAGAGGAG 720

QY 721 CCCAGACAGGGGNGTGAATGAGACACTGCTGCTGGGGGCTCAAGTGGCGGTG 780
DB 721 CCCAGACAGGGGNGTGAATGAGACACTGCTGCTGGGGGCTCAAGTGGCGGTG 780
QY 781 CTGTGAACATCTACTCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 781 CTGTGAACATCTACTCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 CAGTGTCCGTTGGAGAGGCGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 841 CAGTGTCCGTTGGAGAGGCGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 TGAACATGAG 960
DB 901 TGAACATGAG 960
QY 961 GGGTTATAG 1020
DB 961 GGGTTATAG 1020
QY 1021 TGAAGTGTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1021 TGAAGTGTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 TGAAGTGTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1081 TGAAGTGTGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 CCAATTTGATGAG 1200
DB 1141 CCAATTTGATGAG 1200
QY 1201 TGTGTGCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1201 TGTGTGCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 TGTGTGCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 TGTGTGCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 GGGTGTAAAGAGTAAATGAAGTCAATCCGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1321 GGGTGTAAAGAGTAAATGAAGTCAATCCGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 ATTAAGATGATGAG 1440
DB 1381 ATTAAGATGATGAG 1440
QY 1441 ATTAAGATGATGAG 1500
DB 1441 ATTAAGATGATGAG 1500
QY 1501 GGGAG 1560
DB 1501 GGGAG 1560
QY 1561 TCCGTGCAAG 1620
DB 1561 TCCGTGCAAG 1620
QY 1621 CCATCAG 1680
DB 1621 CCATCAG 1680
QY 1681 GATACAG 1740
DB 1681 GATACAG 1740
QY 1741 ATCTGCATCATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
DB 1741 ATCTGCATCATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 1801 GACTCAAAAAAATGGCGCATTCACCAAGTAAATATTATAGAGATAC 1844

DB 1801 GACTCAACAAAAATGGCCATTCAAACCTAATTATACAGATAC 1844

RESULT 2
ABV23195
ID ABV23195 standard; cDNA; 3634 BP.
XX
AC ABV23195;
XX
XX 16-SEP-2002. (first entry)
XX
DE Human prostate expression marker cDNA 23186.
XX
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene: ss.
XX
XX Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PE 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R; Endege WO, Monahan JE.
XX
DR WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim 1; Page 4168-4169; 11750p; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a patient;
XX (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 3634 BP; 876 A; 966 C; 869 G; 918 T; 5 other;
XX
Query Match 99.8%; Score 1840.8; DB 23; Length 3634;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 121 CGCAGCTGACGACGAGGGGGCCGAGAGAGGTGGGCGATCCGCTGTCGAGAGGGCCGCGC 180
DB 191 CGCAGCTGACGACGAGGGGGCCGAGAGAGGTGGGCGATCCGCTGTCGAGAGGGCCGCGC 250
QY 181 GGGCAGCGGGGGGGGGCCGAGAGGGGGGAAAGAGGGGGGGGGGGTCAAGCCGCTGGCC 240
DB 251 GGGCAGCGGGGGGGGGCCGAGAGGGGGGAAAGAGGGGGGGGGGGTCAAGCCGCTGGCC 310
QY 241 GGGCGGGGGGGGGGATGCTGATGCTGAGCGGATGCGGCTGCGGGGGGGGGAGGAGG 300
DB 311 GGGCGGGGGGGGATGCTGATGCTGAGCGGATGCGGCTGCGGGGGGGGGAGGAGG 370
QY 301 TGAAGCAGGCGCAAGAGATGAGAGCGCCGAGAGATGCTGCTTCATGCGCCACGCTACCA 360
DB 371 TGAAGCAGGCGCAAGAGATGAGAGCGCCGAGAGATGCTGCTTCATGCGCCACGCTACCA 430
QY 361 AGGCGCCCAAGAGCAAAATCCAGTTGCTGATGATGATGAGAGTTACCAATTCGCCA 420
DB 431 AGGCGCCCAAGAGCAAAATCCAGTTGCTGATGATGATGAGAGTTACCAATTCGCCA 490
QY 421 CCAAACTGGCGGAGATCTTTGCTGCTGCTGATCTGACAGTCTCCACTGACAGCTACA 480
DB 491 CCAAACTGGCGGAGATCTTTGCTGCTGCTGATCTGACAGTCTCCACTGACAGCTACA 550
QY 481 GTTCAGCTGATCCTACACAGATAGCTGTGATGAGTTTCTCCCGAGAGAGAGC 540
DB 551 GTTCAGCTGATCCTACACAGATAGCTGTGATGAGTTTCTCCCGAGAGAGAGC 610
QY 541 AAACCACTCCAGAGGCGAGCAGCAATTTCTGTGAGAAACATCAGCAGGAGCAATTTG 600
DB 611 AAACCACTCCAGAGGCGAGCAGCAATTTCTGTGAGAAACATCAGCAGGAGCAATTTG 670
QY 601 GAGCGGGGAGATGAGATGACAGAGAGCAATGCTGCTGATTCCTACTGAGAAAC 660
DB 671 GAGCGGGGAGATGAGATGACAGAGAGCAATGCTGCTGATTCCTACTGAGAAAC 730
QY 661 GTGCTCAGGGGAGAGACCCCTTGCTGTGCTAAATATGAGGCTGACACATCAG 720
DB 731 GTGCTCAGGGGAGAGACCCCTTGCTGTGCTAAATATGAGGCTGACACATCAG 790
QY 721 CCCAGACAGGGGTTTATGAGACACTCTGTGCTGAGGGGCTCACTGCGCTGTG 780
DB 791 CCCAGACAGGGGTTTATGAGACACTCTGTGCTGAGGGGCTCACTGCGCTGTG 850
QY 781 CTTTACATCTACTCACTCACTAGATGAGATGAGTGTGAGCAGCTGAGGCTGAGATTG 840
DB 851 CTTTACATCTACTCACTCACTAGATGAGATGAGTGTGAGCAGCTGAGGCTGAGATTG 910
QY 841 CAGTGTGCTTGGAGAGGGGAGTCAAGATGACTTCTGTGTGTATTGACCGCTGTG 900
DB 911 CAGTGTGCTTGGAGAGGGGAGTCAAGATGACTTCTGTGTGTATTGACCGCTGTG 970
QY 901 TGAACATGATGGTGGCAGGCGCAACATGATCCCTGATGATGGGGAGACTTAACCACT 960
DB 971 TGAACATGATGGTGGCAGGCGCAACATGATCCCTGATGATGGGGAGACTTAACCACT 1030
QY 961 GGGTTTAAAGATATCAAAAGCTTTTAAAGATGCCAGGATGTTGGAGAGAGAGCG 1020
DB 1031 GGGTTTAAAGATATCAAAAGCTTTTAAAGATGCCAGGATGTTGGAGAGAGAGCG 1090
QY 1021 TGAATGTGTACAGGCTGTATCAGCTCTCCAAAGCTGGGAGACTGTGTTCCGCGCA 1080
DB 1091 TGAATGTGTACAGGCTGTATCAGCTCTCCAAAGCTGGGAGACTGTGTTCCGCGCA 1150
QY 1081 TGAACGTCAATGATCTGTATACCAAGAGAGTTGATTAATCTTACTGCTGCGGAGAT 1140
DB 1151 TGAACGTCAATGATCTGTATACCAAGAGAGTTGATTAATCTTACTGCTGCGGAGAT 1210
QY 1141 CCATTTTGGATGGCTGAGAGGAGCAGATGATGATTTGGTGGGAAACAAGTGGTG 1200
DB 1211 CCATTTTGGATGGCTGAGAGGAGCAGATGATGATTTGGTGGGAAACAAGTGGTG 1270

QY	1681	GATPACAAGCAGATGTGTACTTGGCTTCCCTAAGAAATGATGAAATACGTCGACGCTTGC	1740
QY	1751	GATPACAAGCAGATGTGTACTTGGCTTCCCTAAGAAATGATGAAATACGTCGACGCTTGC	1810
Db	1741	ATCTGCCATCATTTGATGATGCCACCTTACAGAGCTGACAGATGACCAACAAATATCTGG	1800
QY	1811	ATCTGCCATCATTTGATGATGCCACCTTACAGAGCTGACAGATGACCAACAAATATCTGG	1870
QY	1801	GACTCAACAAATAATGGGCGCTTCAACCCATTAATATTATACAGATAC	1844
Db	1871	GACTCAACAAATAATGGGCGCTTCAACCCATTAATATTATACAGATAC	1914
XX	AAV73924	standard; DNA; 3616 BP.	
XX	AAV73924	(first entry)	
XX	Human SAHH DNA #1.		
XX	S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;		
XX	treatment; infection; cancer; autoimmune disease; detection; diagnosis;		
XX	gene mapping; antisense; therapy; antagonist; immunosays; ss		
XX	Homio sapiens.		
XX	key	Location/Qualifiers	
XX	CDS	56..1558	
XX	FT	/*tag- a	
XX	FT	/product- "SAHH"	
XX	US5854023A		
XX	29-DEC-1998.		
XX	17-JUL-1997;	97US-0896005.	
XX	17-JUL-1997;	97US-0896005.	
XX	(INCY-)	INCYTE PHARM INC.	
XX	Corley NC, Hillman JL, Lal P, Shah P;		
XX	WPI: 1999-094906/08.		
XX	P-PSDB; AAM90061.		
XX	Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase		
XX	for production of recombinant enzyme, useful for diagnosis,		
XX	treatment and prevention of cancers, infections and autoimmune		
XX	diseases		
XX	Disclosure; Fig 1A-I; 40pp; English.		
XX	This sequence encodes a human S-adenosyl-5-homocysteine hydrolase (SAHH).		
XX	The SAHH protein can be used to generate specific antibodies and in drug		
XX	screening to identify specific binding agents. Antagonists of the		
XX	protein are used to treat or prevent a wide range of viral, bacterial,		
XX	fungal, parasitic, protozoal or helminthic infections, many cancers		
XX	(leukemia, lymphoma or solid tumors), and many autoimmune diseases		
XX	(e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes		
XX	mellitus, multiple sclerosis etc). All these conditions may be treated by		
XX	expressing antisense sequences, triplex-forming agents or ribozymes		
XX	directed against the nucleic acid. The nucleic acid and its fragments can		
XX	be used as probes or primers for detecting and quantifying gene		
XX	expression, for diagnosis or monitoring of disease, to identify genetic		
XX	variations, mutations or polymorphisms, in gene mapping and as antisense		
XX	therapeutics. Antibodies are used directly as antagonists, indirectly to		
XX	deliver active agents to SAHH-expressing cells, to diagnose and monitor		
XX	diseases in standard immunoassays, in competitive drug screens and to		
XX	isolate the protein from natural sources.		

PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS Claim 1; SEQ ID NO 5561; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB157137-AB12072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp://wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 1566 BP; 360 A; 438 C; 481 G; 287 T; 0 other;
 SO
 Query Match 37.6%; Score 694.2; DB 23; Length 1566;
 Best Local Similarity 70.4%; Pred. No. 2.9e-175;
 Matches 943; Conservative 0; Mismatches 393; Indels 3; Gaps 1;
 QY 509 TGAATGATGAGTTTCCTCCCGAAGAAAGCAACCACTCCAGAGGCGACGACATTT 568
 DB 225 TGGCGACGAGTCCCGCCGAGGAGAGGTCCAAAGATCCAAAGGCGCTCCGATT 284
 QY 569 CTGTGGAAGACAT---CAAGCAGGACGAATTTGACGCGGAGATGATGATGAGA 625
 DB 285 CTGTGGAAGACATGAGGCGCCGACGATGCTTTGGGCGAGGAGATTCGAGATTCGCGA 344
 QY 626 GCAAGACATGCTGCTGTGATTCACACAGAAAGCTGCTCAGGAGGAGAAAGCCCTTGGC 685
 DB 345 GCGAGAGATCCGCGGATCATATAGCGCTGAAGAGCGGCGGAGCAACAGCTCTGAA 404
 QY 686 TGTGCTAAATATGTGGGCTGTACACATCACAGCCGACAGCGGTGTGATGAGAC 745
 DB 405 GGTATGCCAATGCTGTGGTGTGACACATCAAGCCCACTGCTGCTCATCGAGAC 464
 QY 746 ACCTGTGCGCTGGGGGCTCAGTGCCGCTGCTGTTAATCACTACTACATCAGAA 805
 DB 465 GCTGTGTGAATCTGGGCGCCAGCGCTGCGCTGGCGCTGCAACATCTACTCCAGCA 524
 QY 806 TGAATGATGACACACTGCTGAGCTGAGCTGAGTTCAGTGTGTTGGAAGGCGAGTC 865
 DB 525 CGAAGTGGCTGCTGTGGCCGAGATCCGATCTTTCGCGGCGCGAGAC 584
 QY 866 AGAAGATGACTTGTGTGTATGACCGCTGTGTGAACATGAGTGGTGGACAGCCAA 925
 DB 585 GGAAGAGACTTCTGTGTGTATGATCGCTGTTAAGCGCCGAGAACTGGCAGCCAAA 644
 QY 926 CATGATCTCTGATGATGAGGAGACCTTAACCACTGGGTTTATGAAGATTCCAAAGCT 985
 DB 645 CATGATCTCTGATGATGAGGAGACCTTAACCACTGGGTTTATGAAGATTCCAAAGCT 704
 QY 986 GTTTAAGAAATCCGAGGCAATTTGGAAGAGAGCGTACTGTCTTCAACAGGCTGATCA 1045
 DB 705 GTTCAAGTGTGCAAGGAGATTCGCGAGAGAGCGTCAACCGGAGTGCACCTCTCTATCA 764
 QY 1046 GCTCTCCAAAGCTGGAGGCTGTGTTCGCGCATGAGCTCAATGATTCCTTACCAA 1105
 DB 765 GCTGTCAAGGCGCGCAAGTGAACAGTCCGAGTGAATGATCAACATTCGGTGAACAA 824
 QY 1106 ACAGAAATTTGATTAATCTGCTGCTGCGAGATCCATTTTGGATGGCCCTGAAGAGAC 1165
 DB 825 GACCAATTTTGAACAACCTTACACCTGCAAGGAATCCATCTGGAACGCTCAAGCGCTC 884
 QY 1166 CACAGATGATGATTTTGGGAGAAACAAGTGTGTGATGGCTATGATGATGAGCA 1225
 DB 885 CACGAGATGATGATTTTGGGAGAAACAAGTGTGTGATGGCTATGATGATGAGCA 944
 QY 1226 GGGCTGCTGTGCTGCTCTCAAAAGCTTGTGAACAATGTCTTACATTCGAAATGACCC 1285

DB 945 GGGATGCTCAGGCTCTTAAGGGGAGGCGCTGCTGATGTGTATCATCAGAGATCGATCC 1004
 QY 1286 CATCTGTGCTCTCAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1345
 DB 1005 CATTTGGCGCTCAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1064
 QY 1346 CCGGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1405
 DB 1065 CCGGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1124
 QY 1406 CTTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1465
 DB 1125 CATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1184
 QY 1466 CGATGTGACCAAGCTCCGACCTCCGACCTCCGACCTCCGACCTCCGACCTCCGAC 1525
 DB 1185 CGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1244
 QY 1526 CCATGTATCTGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1585
 DB 1245 TCAATATATCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1304
 QY 1586 TTTGAGCTGCTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAG 1645
 DB 1305 TCTGAGCTGCTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAG 1364
 QY 1646 GGCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1705
 DB 1365 GGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1424
 QY 1706 TCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1765
 DB 1425 GCCCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1484
 QY 1766 TACAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAACAAATATGGCATTC 1825
 DB 1485 GACGAGAGCTGACAGATGACCAAGCAAAATATCTGGAGCTCAACAAATATGGCATTC 1544
 QY 1826 ACCTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1844
 DB 1545 GCCCACTACTAGCAGCTAC 1563
 RESULT 8
 ID ABL29787 standard; DNA; 1776 BP.
 XX ABL29787;
 AC 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 40834.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 OS Drosophila melanogaster.
 XX
 PN W0200171042-A2.
 XX
 PD 27-SEP-2001;
 XX
 PE 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EM.
 XX
 DR WPI; 2001-656860/75.

XX Venter, JC, Adams M, LI PMD, Myers EW;
 XX WPI; 2001-656860/75.
 DR P-PSDB; ABB59589.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 5558; 21bp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB101875) and the encoded proteins
 CC (AB101840-AB101875).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp://wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 4406 BP; 1112 A; 1039 C; 1064 G; 1191 T; 0 other;
 SQ
 Query Match 26.4%; Score 486.2; DB 23; Length 4406;
 Best Local Similarity 66.9%; Pred. No. 2.1e-119;
 Matches 742; Conservative 0; Mismatches 303; Indels 64; Gaps 1;

QY 1396 CACGGAGACACTGGATCGCATGAAAAACAGTTGTATCGTATGCAATATGGCCACTCCA 1455
 DB 1452 TGGCGAGACATGACAAAGATGAGTGGTCATCGTATGCAATATGGCCACTCCA 1393
 QY 1456 ACACAGAAATCGATGTGACACAGCTCCGACCTCCGAGCTGAGAGTGGACAGTACGT 1515
 DB 1392 ACACAGAAATCGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1333
 QY 1516 CTCAGGTGACATGTCATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1575
 DB 1332 CCCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1273
 QY 1576 GTCTACTCAATTTGAGTGTCTCCACAGTTCACACTTTGTCCTCCATCAGACCCCAA 1635
 DB 1272 GGTGTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1213
 QY 1636 CACAGGTTTGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1695
 DB 1212 CCCAGGCTTGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1153
 QY 1696 TGTACTTGTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1755
 DB 1152 TCTACTTGTCTCTCTTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1093
 QY 1756 ATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACCTCAGCAAAATATG 1815
 DB 1092 ATGCCACCTTACAGAGCTGACAGATGACCAAGCAAAATATCTGGACCTCAGCAAAATATG 1033
 QY 1816 GGCCTTCAACCTTAATTTATTCAGATAC 1844
 DB 1032 GTCCTTCAAGCCCACTACTACCGCTAC 1004

RESULT 10
 ABL79045
 ID ABL79045 standard; cDNA; 553 BP.
 XX
 AC ABL79045;
 XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Human ovarian cancer related cDNA clone SEQ ID NO:2023.
 XX
 KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200192581-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001MO-US17756.
 XX
 PR 26-MAY-2000; 2000US-207484P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Algate PA, Harlocker SL, Jones R;
 XX
 DR WPI; 2002-122075/16.
 XX
 PT Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PT polypeptide -
 XX
 PS Claim 1; SEQ ID 2023; 489bp; English.
 CC The present invention describes a composition (i) comprising: carriers
 CC and immunostimulants; and a polypeptide (ii) of a ovarian tumour
 CC polypeptide encoded by a polynucleotide (iii) having a cDNA sequence
 CC (SI) from the 10912 nucleotide sequences as given in ABL77023 to

CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 CC population of (II), or antigen presenting cells that express (II).
 CC (1) has cytostatic activity. An oligonucleotide (IV) that hybridizes to
 CC (S1) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridizing to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridizing to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC (III) and/or (II) is useful for stimulating and/or expanding T cells
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.

SO Sequence 553 BP; 159 A; 136 C; 128 G; 127 T; 3 other;

Query Match 25.0%; Score 461.4; DB 24; Length 553;
 Best Local Similarity 97.3%; Pred. No. 3.3e-113;
 Matches 510; Conservative 0; Mismatches 9; Indels 5; Gaps 4;

DB 1322 GGTGGTAACTTAATTAAGTATCGCGCAAGTCGATGTAATACTTGACAGAGAAA 1381
 |||||||
 DB 1 GGTGGTAACTTAATTAAGTATCGCGCAAGTCGATGTAATACTTGACAGAGAAA 60
 |||||||
 OY 1382 TAAGATGTGTGACAGGAGGACCTGGATTCGATTAATACTTGACAGAGAAA 1441
 |||||||
 DB 61 TAAAGATGTGTGACAGGAGGACCTGGATTCGATTAATACTTGACAGAGAAA 120
 |||||||
 OY 1442 TATGGGCACTCCACACAGAAATCGATGTGACAGGCTCCGACCTCGAGAGTGCAGTG 1501
 |||||||
 DB 121 TATGGGCACTCCACACAGAAATCGATGTGACAGGCTCCGACCTCGAGAGTGCAGTG 180
 |||||||
 OY 1502 GGAGCGATGTCGTCAGTGTGACATGTCATCTGGCCAGATGCGAAGAGAGTGTGCT 1561
 |||||||
 DB 181 GGAGCGATGTCGTCAGTGTGACATGTCATCTGGCCAGATGCGAAGAGAGTGTGCT 239
 |||||||
 OY 1562 CTTGGCAGAGGTCGTCAGTGTGACATGTCATCTGGCCAGATGCGAAGAGAGTGTGCT 1621
 |||||||
 DB 240 CTTGGCAGAGGTCGTCAGTGTGACATGTCATCTGGCCAGATGCGAAGAGAGTGTGCT 299
 |||||||
 OY 1622 CATCAGAGCCACACAGAGGCTTGGCAGTGTGATGAATCTATATGACCCGAGGGCG 1681
 |||||||
 DB 300 CATCAGAGCCACACAGAGGCTTGGCAGTGTGATGAATCTATATGACCCGAGGGCG 359
 |||||||
 OY 1682 ATACAGAGGATGTGTCCTCTCTTAAGAAA-TGGATGATAGCTTGCAGCTTGC 1740
 |||||||
 DB 360 ATACAGAGGATGTGTCCTCTCTTAAGAAA-TGGATGATAGCTTGCAGCTTGC 419
 |||||||
 OY 1741 ATCTGCATCATTTGATGCCACCTTACAGAGTGCAGATGACCAAGCAAAATATCTGG 1800
 |||||||
 DB 420 ATCTGCATCATTTGATGCCACCTTACAGAGTGCAGATGACCAAGCAAAATATCTGG 478
 |||||||
 OY 1801 GACTCAACAAAATGGCCATTCAAACCTTAATTATACAGATAC 1844
 |||||||
 DB 479 GACTCAACAAA--TGGCCATTCAAACCTTAATTATACAGATAC 520
 |||||||

RESULT 11
 ABV44141/C
 ID ABV44141 standard; cDNA; 504 BP.

XX AC ABV44141;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 44132.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.
 XX PN WC0200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-0505171.

PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

DR WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer

PS Claim 1; Page 8770; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

SO Sequence 504 BP; 117 A; 150 C; 116 G; 121 T; 0 other;

Query Match 22.9%; Score 422; DB 23; Length 504;
 Best Local Similarity 100.0%; Pred. No. 1.1e-102;
 Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 706 GTACACATCATCAGACCCACAGCGGTGTGATGAGACATCTGTGCCCTGGGGCTC 765
 |||||||
 DB 502 GTACACATCATCAGACCCACAGCGGTGTGATGAGACATCTGTGCCCTGGGGCTC 443
 |||||||
 OY 766 AGTCCGCTGTGTGTTGTAACATCTACTCACTCGAATGAAGTGTGACAGACTGG 825
 |||||||
 DB 442 AGTCCGCTGTGTGTTGTAACATCTACTCACTCGAATGAAGTGTGACAGACTGG 363
 |||||||
 OY 826 CTGAGCTGAGTGTGAGTGTGCTTGGAGGGCGAGTGAAGATGACTTGTGTGT 885
 |||||||
 DB 382 CTGAGCTGAGTGTGAGTGTGCTTGGAGGGCGAGTGAAGATGACTTGTGTGT 323
 |||||||
 OY 886 GTATTAGCCGCTGTGATGACATGATGAGTGTGACAGCCACATGATCTTGATGATGGG 945
 |||||||
 DB 322 GTATTAGCCGCTGTGATGACATGATGAGTGTGACAGCCACATGATCTTGATGATGGG 263
 |||||||
 OY 946 GAGACTTAACCCACTGGGTTTATAGAGATATCCAAACGTTTAAAGATCCGAGGCA 1005
 |||||||
 DB 262 GAGACTTAACCCACTGGGTTTATAGAGATATCCAAACGTTTAAAGATCCGAGGCA 203
 |||||||
 OY 1006 TTGTGAGAGAGAGCGTGAAGTGTGTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC 1065
 |||||||
 DB 202 TTGTGAGAGAGAGCGTGAAGTGTGTCACAGGCTGTATCAGCTCTCCAAAGCTGGGAAGC 143
 |||||||

QY 1066 TCTGTGTCGGCCATGACATGATCTCTTACCAAGAAAGTTGATCTGT 1125
DB 1126 AC 1127
QY 82 AC 81

RESULT 12
ABL29786
ID ABL29786 standard; DNA; 5537 BP.
AC ABL29786;
DE 26-MAR-2002 (first entry)
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 40831.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; de.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PERE) PE CORP NY.
XX Venter JC, Adams M, Li PHD, Myers EW;
XX WPI; 2001-656860/75.
XX

DR WPI; 2001-656860/75.
XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX

PS Claim 1; SEQ ID NO 40831; 21pp + Sequence Listing; English.
XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB557737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC

CC Sequence 5537 BP; 1632 A; 1054 C; 1158 G; 1693 T; 0 other;
XX

Query Match: 22.9%; Score 421.6; DB 23; Length 5537;
Best Local Similarity 59.7%; Pred. No. 4,4e-102;
Matches 852; Conservative 0; Mismatches 439; Indels 137; Gaps 3;

QY 438 TCTGTGTCGGCCATGACATGATCTCTTACCAAGAAAGTTGATCTGT 497
DB 2826 TCTGTGTCGGCCATGACATGATCTCTTACCAAGAAAGTTGATCTGT 2885

QY 498 ACAGATAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 557
DB 2886 GGAAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2945

QY 558 AGCAGCAATTTCTGTGTGAAGAAATCAAGCAGCAGATTTGGACCCGGAGATTGAG 617

DB 2946 GGCACATGATCTTTGGCTGGAAGAGATCTCGAAGAGCTTTGGAGAGGAGATCGAG 3005
QY 618 ATGACAGCAAGACATGATCTGCTGATTTACCTAGGAAAGCTGCTGAGGAGAG 677
DB 3006 ATGACAGCAAGATGCTGAGGAGATGATGATGATGATGATGATGATGATG 3065

QY 678 CCCTTGGCTGCTGATTAATAGTGGCTGTACACATACAGACCCAGACGGGTGG 737
DB 3066 CCCCTAAAGGCTGCAATATGCTGAGTACACACAGCTCAATGCTGAGCGAGCTG 3125

QY 738 ATTGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 797
DB 3126 ATGAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3185

QY 798 ACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 857
DB 3186 ACACAAAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3245

QY 858 GCGAGATCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 917
DB 3246 GAGAGACGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3305

QY 918 CAGGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 977
DB 3306 CAGGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3365

QY 978 CCAAGCTGTTTAAAGATCCGAGGATTTGGAAGAGAGCTGAGCTGCTGCTGCTG 1037
DB 3366 CCGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3425

QY 1038 CTGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1097
DB 3426 CTGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3485

QY 1098 GTTACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1150
DB 3486 GTTACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 3545

QY 1151 ----- 1150
DB 3546 GCATAGTACTGATGATTAAGAAAGTACTATCTGATGCTGCTGCTGCTGCTG 3605

QY 1151 -----TGCCCTGAAGAGAGACACAGATGATGATGATGATGATGATGATG 1206
DB 3606 CTTCAGATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3665

QY 1207 GCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1266
DB 3666 GCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3725

QY 1267 ACATTAACCAAG 1326
DB 3726 ATGTTACGAG 3785

QY 1327 TAAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1386
DB 3786 TAAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3845

QY 1387 ATGTTAG 1446
DB 3846 ATGTTAG 3905

QY 1447 GGCATCCCAACAG 1471
DB 3906 GACATTCCTGCTGAGAGATGATGATGATGATGATGATGATGATGATGATG 3965

QY 1472 -----GACAGCTCCCATCTCGGAGCTGAGCGG 1503
DB 3966 TATATCAAAAGGTTTAAACCACTTTAGATGATGATGATGATGATGATGATG 4025

QY 1504 AGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1563

Db 4026 AGCGTCCGTTCTCAAGTGCATCATCAGTGGCCGAGCGAGATGATCATTTTGC 4085

QY 1564 TGGCAGAGGTCGTCATCTCAATTTGAGCTGCTCCACAGTCCACCTTTGTTGTCCA 1623

Db 4086 TGGCGAGGAGGAGACTGTGATTTGCTTCACCATTTGCTTGGTGGATCCG 4145

QY 1624 TCACAGCCACACAGAGCTTTGGACATGATAGAACTATATATACCCAGGCGCAT 1683

Db 4146 TGGCTCATCCACAGGCTTTGGCCCTGATTTGAATCTTCTCAGCGCC---AGGAGAT 4202

QY 1684 ACAAGCAGAGTGTACTTCTCTTAAAGAAATGATGATGATG 1731

Db 4203 ATAGTGGATGCTACTGCTGCAAGAAATGGATGTTCTTAG 4250

RESULT 13

AAZ80766/c

ID AAZ80766 standard; cDNA; 636 BP.

XX

AC AAZ80766;

XX

DT 07-APR-2000 (first entry)

XX

DE Human colon cancer cell line SW480 cDNA clone SEQ ID NO:850.

XX

KM Human: gene expression product; diagnosis; tumour; colon cancer;

KM colorectal adenocarcinoma; cell line SW480; cell proliferation;

KW cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;

KW hyperplasia; ds.

XX

OS Homo sapiens.

XX

PN W0964576-A2.

XX

PD 16-DEC-1999.

XX

PE 09-JUN-1999; 99MO-IB01062.

XX

PR 10-JUN-1998; 98US-0088801.

XX

PA (FARB) BAYER CORP.

XX

PI Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;

PI Carroll E, Catino TJ, Dertl A, Ford DM, Lewis ME, Monahan JE;

PI Schlegel R;

XX

DR MPI; 2000-087220/07.

XX

PT Novel nucleic acids, used to develop products for the diagnosis and

PT treatment of disorders involving unwanted cell proliferation,

PT particularly cancers, especially colon cancer

XX

PS Claim 15; Page 469; 469pp; English.

XX

XX AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from

CC the human colorectal adenocarcinoma (colon cancer) cell line SW480. The

CC cDNA clones can be used to generate antisense oligonucleotides which

CC can be used for antisense therapy. Methods and products from the present

CC invention can be used for identifying and/or classifying cancerous cells

CC present in a human tumour, particularly in solid tumours, e.g.

CC carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones

CC can be used for developing agents for the diagnosis and treatment of

CC disorders involving unwanted cell proliferation, such as neoplasia,

CC dysplasia or hyperplasia.

XX

XX Sequence 636 BP; 156 A; 166 C; 137 G; 147 T; 30 other;

SO

Query Match 22.0%; Score 406.4; DB 21; Length 636;

Best Local Similarity 99.5%; Pred. No. 1.8e-98;

Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 707 TACACATCACAGCCACACAGC-GGTGTGATTGAGACACTGTGCGGAGGCGTC 765

|||||

Db 420 TACACATCACAGCCACACAGCGGGTGTGATTGAGACACTGTGCTGGGGCGTC 361

QY 766 AGTGGCGGTGCTGCTCTTGTATACATCTACTCAACTCAGATGATAGCTGAGCAGCTGG 825

Db 360 AGTGGCGGTGCTGCTCTTGTATACATCTACTCAACTCAGATGATAGCTGAGCAGCTGG 301

QY 826 CTGAGGCTGAGATTGCACTGTGCTTGGAGGGGAGTCAGATGATGATGATGATGAT 885

Db 300 CTGAGGCTGAGATTGCACTGTGCTTGGAGGGGAGTCAGATGATGATGATGATGAT 241

QY 886 GTATTGACCGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 945

Db 240 GTATTGACCGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 181

QY 946 GAGACTTACCCACAGGGGTTATAGAGTATCCAAAGTGTATAGAGATCCAGGCA 1005

Db 180 GAGACTTACCCACAGGGGTTATAGAGTATCCAAAGTGTATAGAGATCCAGGCA 121

QY 1006 TTGTGGAAGAGAGCGGTGACTGTTTTCACAGGCTGATCAGCTCCAAAGCTGGGAAGC 1065

Db 120 TTGTGGAAGAGAGCGGTGACTGTTTTCACAGGCTGATCAGCTCCAAAGCTGGGAAGC 61

QY 1066 TCTGTGTTCCGCGCATGACGTCAATGATTTGTACCAACAGAGTTGATTAATCTGT 1125

Db 60 TCTGTGTTCCGCGCATGACGTCAATGATTTGTACCAACAGAGTTGATTAATCTGT 1

RESULT 14

AACT7823

ID AACT7823 standard; cDNA; 2200 BP.

XX

AC AACT7823;

XX

DT 08-FEB-2001 (first entry)

XX

DE Human cancer associated gene sequence SEQ ID NO:217.

XX

KM Human: cancer associated gene; cancer antigen; detection; cancer;

KM diagnosis; cytostatic; proliferative; vulnery; immunomodulator;

KM antidiabetic; antisthmatic; antihemetic; antirheumatic; antiviral;

KM antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;

KM dermatological; neuroprotective; thrombolytic; coagulant; nootropic;

KM vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;

KM immune disorder; haematopoietic cell disorder; autoimmune disorder;

KM allergic reaction; graft versus host disease; organ rejection;

KM haemostatic; thrombolytic; cardiovascular disorder; infection;

XX

OS Homo sapiens.

XX

PN W0200055350-A1.

XX

PD 21-SEP-2000.

XX

PE 08-MAR-2000; 2000MO-US05882.

XX

PR 12-MAR-1999; 99US-0124270.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Ruben SM;

XX

DR MPI; 2000-587533/55.

XX

DR P-PSDB; AAB43614.

XX

PT Novel isolated nucleic acids comprising sequences encoding peptides

PT useful for treating or diagnosing e.g. cancer

XX

PS Claim 1; Page 787-788; 2352pp; English.

XX

XX AACT7607 to AACT7848 encode the human cancer associated proteins given

CC in AAB43398 to AAB44239. The proteins can have activities based on the

CC tissues and cells the genes are expressed in. Example of activities

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 23:48:04 ; Search time 82.4279 Seconds
(Without alignments)
6860.687 Million cell updates/sec

Title: US-09-782-051-1_COPY_1_1844

Perfect score: 1844
Sequence: 1 ggcggcgccgagctgcgagct.....aacctattatcacagatc 1844

Scoring table: IDENTITY_NDC
Gapop 10.0 ; Gapext 1.0

Searched: 441367 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA.*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/6C.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/6D.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	406.4	22.0	636	4	US-09-328-111-850
2	314.4	17.0	2211	4	US-09-318-448-26
3	314.4	17.0	2211	4	US-09-347-878-2
4	226.2	12.8	4411529	4	US-09-103-840A-1
5	222.6	11.6	1812	2	US-08-669-536-1
6	214.6	11.6	1767	3	US-08-930-894-1
7	117	6.3	2226	4	US-08-896-005-2
8	117	6.3	2226	4	US-09-347-878-3
9	72.6	3.9	289	1	US-08-204-740-8
10	72.6	3.9	289	3	US-09-081-167A-8
11	72.6	3.9	289	3	US-09-081-167A-8
12	72.6	3.9	289	4	US-09-416-833-8
13	72.6	3.9	289	5	PCT-US95-02521-8
14	68.4	3.7	285	1	US-08-204-740-6
15	68.4	3.7	285	3	US-09-081-167A-6
16	68.4	3.7	285	3	US-09-081-167A-6
17	68.4	3.7	285	4	US-09-416-833-6
18	68.4	3.7	285	5	PCT-US95-02521-6
19	68.2	3.7	584	4	US-08-998-416-249
20	57.4	3.1	7218	1	US-08-233-463-14
21	52.2	2.8	320	4	US-09-165-264-13
22	52.2	2.8	320	4	US-09-165-264-7
23	51.6	2.8	319	4	US-09-165-264-8
24	51.2	2.8	320	4	US-09-165-264-14
25	50.4	2.7	320	4	US-09-165-264-11
26	50.2	2.7	153331	3	US-09-128-155-16
27	48.8	2.6	1926	4	US-09-249-585A-2

28	48.8	2.6	2580	3	US-09-050-863-2	Sequence 2, Appl
29	48.8	2.6	2580	4	US-09-359-081-2	Sequence 2, Appl
30	48.8	2.6	5452	2	US-09-130-114-1	Sequence 1, Appl
31	48.8	2.6	9600	4	US-08-910-647-1	Sequence 1, Appl
32	48.8	2.6	9600	4	US-09-620-925-1	Sequence 1, Appl
33	48.8	2.6	10596	1	US-07-884-811-15	Sequence 15, Appl
34	48.8	2.6	10596	1	US-07-885-971-15	Sequence 15, Appl
35	48.8	2.6	10596	1	US-08-087-783A-15	Sequence 15, Appl
36	48.8	2.6	10596	1	US-08-194-088B-15	Sequence 15, Appl
37	48.8	2.6	10596	2	US-08-194-087-15	Sequence 15, Appl
38	48.8	2.6	10596	5	PCT-US93-04648-15	Sequence 15, Appl
39	48.4	2.6	318	4	US-09-165-264-12	Sequence 12, Appl
40	47.6	2.6	1969	4	US-09-098-628-11	Sequence 1, Appl
41	47.4	2.6	1926	4	US-09-249-585A-4	Sequence 4, Appl
42	47.4	2.6	1931	2	US-09-130-114-2	Sequence 2, Appl
43	46.4	2.5	12001	1	US-08-458-568A-11	Sequence 11, Appl
44	46.2	2.5	405	2	US-08-903-800A-3	Sequence 3, Appl
45	45.6	2.5	8438	1	US-07-945-283-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-328-111-850/C
Sequence 850, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinhilber, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Carino, Theodore J.
APPLICANT: Dertl, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Monahan, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 850
LENGTH: 636
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)...(636)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-850

Query Match 22.0%; Score 406.4; DB 4; Length 636;
Best Local Similarity 99.5%; Pred. No. 6.4e-97;
Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
DB 420 TACACACATACAGCCACAGC-GGTGTGATTGAGACACTGTGCTGGGGCTC 765
DB 420 TACACACATACAGCCACAGC-GGTGTGATTGAGACACTGTGCTGGGGCTC 361
DB 766 AGTCCCGCTGCTGCTTAACTACTCACTAGATGAGTACTGACACTG 825
DB 360 AGTCCCGCTGCTGCTTAACTACTCACTAGATGAGTACTGACACTG 301
DB 826 CAGAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 885

Db 300 CTGAGCTGAGATTGCAAGTGTTCCTTGGAAGGGGAGTCAAGATGACTTCTGTGTGT 241
QY 886 GTATTGACCCCTGTGTGAACATGATGGGTGGCAGGCCAACATATCTGTGATGAGG 945
Db 240 GTATTGACCCCTGTGTGAACATGATGGGTGGCAGGCCAACATATCTGTGATGAGG 181
QY 946 GAGACTTAACCCACTGGGTATTAAGAAATGATCCAAACGCTTTAAAGATCCGAGGCA 1005
Db 180 GAGACTTAACCCACTGGGTATTAAGAAATGATCCAAACGCTTTAAAGATCCGAGGCA 121
QY 1006 TTGTGAAGAGAGCGTACTGTGTTCACAGCGTGTATCAAGCTCTCCAAAGCTGGGAAGC 1065
Db 120 TTGTGAAGAGAGCGTACTGTGTTCACAGCGTGTATCAAGCTCTCCAAAGCTGGGAAGC 61
QY 1066 TCTGTGTTCGGGCGCATATAGCTCAATATCTGTACCAACAGAGTTGATTAAGCTGT 1125
Db 60 TCTGTGTTCGGGCGCATATAGCTCAATATCTGTACCAACAGAGTTGATTAAGCTGT 1

RESULT 2

US-09-318-448-26
; Sequence 26, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stenroos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-26

Query Match 17.0%; Score 314.4; DB 4; Length 2211;
Best Local Similarity 61.9%; Pred. No. 1.2e-72;
Matches 498; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

QY 580 ACATCAAGCAGGAGATTGGACGCCGGAGATTGAGATGCAAGACCAAGATGTCTG 639
Db 79 ACATCGCGCTGGCTGCTGGGAGCGAAGGCCCTGACATGCTGGAAGAGATGCGCG 138
QY 640 CTGTGATTCTACCTCAGAAACGTGCTCAGGGGAGAAAGCCCTGGCTGTGCTAAATAG 699
Db 139 GCGTGTATGCGTATGCGGAGCGGTACTCGGCTCCAAAGCCACTGAAGGGCGCCGATCG 198
QY 700 TGGGCTGTACACATCAGACGCCAGACAGCGGTGTGATTGAGACACTGTGCGCTGG 759
Db 199 CTGGCTGCTGCATGACATGACCGGTGAGAGCGCCCTCCTCATGTGAGACCTGTGACCTGG 258
QY 760 GGGCTAGTGGCGCTGTGTCTGTGAACATCTACTCAACTCAAGATGAAGTGTGACAG 819
Db 259 GTGCTAGTGTGAGTGTGCTGCTGCAACATCTTCTCCAGCCAGAACCAATGCGGGCGCTG 318
QY 820 CACTGCTAGGCTGAGTGTGAGTGTGCTTGTGGAAGGGCGAGTGAAGATGACTTCT 879
Db 319 CCATGCGCAAGGCTGACATTCGGGTGTATGCTCGAAGGGCGGAAGGAGAGAGATAC 378
QY 880 GGTGTGTATGACCGCTGTGTGAACATGATGGGTGGCAGGCCAACATGATGCTGGATG 939
Db 379 TGTGTGTATGAGCAGACCGCTGTGTACTTCAAGAGAGGGGCCCTCAACATGATTTCTGAGC 438
QY 940 ATGAGGAGACTTAACCACTGGGTATTAAGAAATGATCCAAACGCTTTAAAGATGATCC 999
Db 439 ACAGGGGAGACTTCAACCACTGATCACAACAGATACCGGAGCTTGTGCGCAGGATCC 498
QY 1000 GAGGCTTTGGAAGAGAGCGTGTGCTGTGTCACAGGCTGTATCAAGCTCTCCAAAGCTG 1059
Db 60 GAGGCTTTGGAAGAGAGCGTGTGCTGTGTCACAGGCTGTATCAAGCTCTCCAAAGCTG 1

Db 499 GAGCATCTCTGAGAGACACAGACTGGGTCCACACCTCTACAAATGATGATGCCATG 558
QY 1060 GGAAGCTCTGTGTTCGGGCCATGAAACGTCATGATGTTGTTACCAACAGAGTTGATA 1119
Db 559 GGAATCTCAAGGTGCTGTCCATCATGTATGATGATGATGATGATGATGATGATGATG 618
QY 1120 ACTGTACTGTGCGGAGAAATCATTTTGTGATGGCTGTGAAGAGACACAGATGATGAT 1179
Db 619 ACCTGTATGCTGTGCGGAGAGTCCCTCATAGATGATGATGATGATGATGATGATGATG 678
QY 1180 TTGTGGGAAGAACAGT 1239
Db 679 TTGTGGGAAGAACAGT 738
QY 1240 CTCTCAAGCTCTGTGAGCAATGTCTTACATTAACGAATGACCCCATGCTGTGTGTC 1299
Db 739 CCGTGGCGGGGTGTGTGAGCCCGCTCATCATCAGCAGATGATGATGATGATGATGATG 798
QY 1300 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
Db 799 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 858
QY 1360 TCGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1383
Db 859 TCTTGTACACACACAGGCTGTA 882

RESULT 3

US-09-347-878-2
; Sequence 2, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25865-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human S-adenosylhomocysteine hydrolase cDNA
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: M61831/GenBank
US-09-347-878-2

Query Match 17.0%; Score 314.4; DB 4; Length 2211;
Best Local Similarity 61.9%; Pred. No. 1.2e-72;
Matches 498; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

QY 580 ACATCAAGCAGGAGATTGGACGCCGGAGATTGAGATGCAAGACCAAGATGTCTG 639
Db 79 ACATCGCGCTGGCTGCTGGGAGCGAAGGCCCTGACATGCTGGAAGAGATGCGCG 138
QY 640 CTGTGATTCTACCTCAGAAACGTGCTCAGGGGAGAAAGCCCTGGCTGTGCTAAATAG 699
Db 139 GCGTGTATGCGTATGCGGAGCGGTACTCGGCTCCAAAGCCACTGAAGGGCGCCGATCG 198
QY 700 TGGGCTGTACACATCAGACGCCAGACAGCGGTGTGATTGAGACACTGTGCGCTGG 759
Db 199 CTGGCTGCTGCATGACATGACCGGTGAGAGCGCCCTCCTCATGTGAGACCTGTGACCTGG 258
QY 760 GGGCTAGTGGCGCTGTGTCTGTGAACATCTACTCAACTCAAGATGAAGTGTGACAG 819
Db 259 GTGCTAGTGTGAGTGTGCTGCTGCAACATCTTCTCCAGCCAGAACCAATGCGGGCGCTG 318
QY 820 CACTGCTAGGCTGAGTGTGAGTGTGCTTGTGGAAGGGCGAGTGAAGATGACTTCT 879
Db 319 CCATGCGCAAGGCTGACATTCGGGTGTATGCTCGAAGGGCGGAAGGAGAGAGATAC 378

OY	880	GGTGGTGTATTTGACCCGCTGTGTATACATGGATGGGGGACGACCAATGATCCTGGATG	939
Db	379	TGTGTGTCATTGACACGACCCCTTACTTCAAGGACGGGCCCTCCATCATGATTTCTGGACG	438
OY	940	ATGGGGGAGACTTAAACCACATGGGTTTATTAAGAAATATCCAAACGTGTTTAAGAAGATCC	999
Db	439	ACGGGGGGGACCTCACCACCACTTATCCACACCAAGTAAACCCGACGCTTCTGCCAGGCATCC	498
OY	1000	GAGGCACTTGTGGAAGAGAGCGTACGTAGTGTTCACAGGCTGTATCAGCTCTCCAAAGCTG	1059
Db	499	GAGGCACTCTGTGAGAGAACACAGACAGTGGGGTCCACAACCTGTAAAGAAATGATGGCCATG	558
OY	1060	GGAAGCTCTGTGTTCGCGGCATATACGTCATAGTATCTGTTCACCAACGAAGTTTGATA	1119
Db	559	GGAATCTTCACAGGTGCTCGCATCAATGTCATATACCTCCCTCACCAAGACAGTTTGGACA	618
OY	1120	ACCTGTACTGCTGGCCGGAATTCATTTTGGATGGCCGTGAAGAGGACACAGATGTGATGT	1179
Db	619	ACCTCTATGGGTGGCGGGAATCCCTCATATGATGGCATCAAGGGGCGACAGATGTGATGA	678
OY	1180	TTTGCTGGGAAAACAAGTG	1239
Db	679	TTGGCGGCAAGGTAGCGGTGTGTGACAGGCTATGTGTATGTGGGCAAGGGCTGTGGCCAGG	738
OY	1240	CTCTCAAAAGCTCTTGAGACAAATTGTCTACATTAACCAATTCGACCCCATCTGTGCTTGC	1299
Db	739	CCCTGCGGGGGTTTGGAGCCCGGCTCATCTACACGAGATTGACCCCATCAACGCGCATGC	798
OY	1300	AGGCGCTGCATGGAAGGGTTTCAAGGTTGGTAAAGCTAAATGAAGTCAATCCGCGCAATGTGATG	1359
Db	799	AGGCTGCCATGGAAGGGCTATAGAGGTGACCAACATGATGAGGCTCTGTCTCAGGAGGACCA	858
OY	1360	TCGTAATATCTTCGACAGAAATA	1383
Db	859	TCTTTGTACACACACACAGGCTGTGA	882

RESULT 4
 US-09-103-840A-1/c
 : Sequence 1, Application US/09103840A
 : Patent No. 6294328
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: FLEISCHMAN, Robert D.
 :
 : APPLICANT: WHITE, Owen R.
 : APPLICANT: FRASER, Claire M.
 : APPLICANT: VENTER, John C.
 :
 : TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 :
 : TITLE OF INVENTION: TUBERCULOSIS
 : FILE REFERENCE: 2436-2007.00
 : CURRENT APPLICATION NUMBER: US/09/103,840A
 : CURRENT FILING DATE: 1998-06-24

```

; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

[illegible]

Db 3628916 ACGAGCTGGGTGACCAAGTCCAAATTGCACACAAAGTACCGGCACTCGGCCTTCCCTGATTCG 3628857

OY 1150 ATGGCCTTAAGAGGACCCACAGATGTGATGTGTGGGGGAAACAAAGTGGTGGTGTGGCT 1209

Db 3628856 ACGGCATCAACCGGGGACCCGACCGACCGGTGATGGCGGGTAAAGAGTCTCTCATCTGCGGGCT 3628797

OY 1210 ATGTGAGGTAGGGCAAGGGCTGCTGTGCTCTCAAAAGCTTCTGGACCAATTGTCAC 1269

Db 3628796 ACGGCAGACGTGCGTAAGGGCGTCTCGGAGGGCGAATGAAGGGCCAGAGCGCGGGCTTCGC 3628737

OY 1270 TTACCGAATATGACCCCATCTGTGTCTGCGAGCGCTGCATGGATGGATGGTTCAGGGGTGA 1329

Db 3628736 TCACCGAGATCTGACCCCATCAACCGGTGACGGCATATGAGAGGGCTTCGACCTGGTCA 3628677

OY 1330 AGCTAAATGAAGTCAATCCGGCAAGTCGATGTGCTATATACCTTGACACAGAAATGAATG 1389

Db 3628676 CCGTCTGAGAGAGCCATGGGGAGCGCCGACACATCTGTGTATACCGGACCGGGACAAAGACA 3628617

OY 1390 TAGTCACACAGGAGACATTGGATCGCATGAAAAACAGTTGATCGTATGCAATATGGGCC 1449

Db 3628616 TCATCATGCTCTGACACACATTAAAGCGATGAAGAGCACACCGCATCTGGAAATATGGGCC 3628557

OY 1450 ACTCCACACAGAAATTCATGTGTGACACACCTCCGACATCCGGAGCTGAGCGTGAAGAGCGAG 1509

Db 3628556 ACTTCGACACAGATATGACATGGCCGGGCTTGAGACCGCTCCGGGGCGACAGGGTCAACG 3628497

OY 1510 TACGTTCTCAGGTGGACCATGTATCTGGCCAGA--TGGCAACAGAGTGTGCTCTCTGG 1566

Db 3628496 TCAGAGCCTCAGGTGACGACTGTGGACCTTTGGCGACACGGGGCGGCTCGATCATCTGTGCT 3628437

OY 1567 CAGAGGGTGTCTACTCAATTTGAGCTGTCTCCACAGTTCCACC---TTTGTCTGTCTCA 1623

Db 3628436 CCGAGGGGGCGGCTCTGAACCTGGGCAATGCGACCGGGCACCCCTCGTCTGTGTAGACA 3628377

OY 1624 TCACAGCCACACACAGAGGTTTGGGACATGATGTGAAGCTATATATGCACCCGAGGGGCAT 1683

Db 3628376 ACAGCTTCGCTACCCACACAGATCCGCCAGATCGAGCTGTGACCAAGACGACAG---T 3628320

OY 1684 ACAAGCAGAGTGTGTACTGTTCTTCTTAAGAAATGGATGAAATACGTTCGCCAGCTTGACATC 1743

Db 3628319 ACGACAAAGAGGTGTACCGGGCTGGCCAAAGCACTCTGACAGAGAAAGTGGCTGAAATCATG 3628260

OY 1744 TGGCATCATTTGATGCCACCTTAACAGAGCTACAGATGACCAACCAAAATATCTGGGAC 1803

Db 3628359 TCGAGGCCCTTGGGGGTACCTCTGACCAAGCTGACCAAGAGACGAGCGCAATACCTCGCGC 3628200

OY 1804 TCAACAAAATGGGCATTCAACCTAATTTATTACAGATAC 1844

Db 3628199 TCGAGCTGAGGGTCTCTACAGACGGGACACACTACCGCTAC 3628159

RESULT 5
US-08-669-536-1
Sequence 1, Application US/08669536
Patent No. 5910444.
GENERAL INFORMATION:
APPLICANT: MASUDA, CHIKARA
APPLICANT: MASUDA, YUKO
APPLICANT: TANAKA, HIDEO
APPLICANT: KUMATA, SHIGERU
TITLE OF INVENTION: ORGANISMS IN WHICH THE EXPRESSION OF
TITLE OF INVENTION: S-DENOSYLHOMOCISTEINE HYDROLASE GENE IS INHIBITED
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/669,536
 FILING DATE:
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: MURPHY JR, GERALD M
 REGISTRATION NUMBER: 28,977
 REFERENCE/DOCKET NUMBER: 1254-128
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1812 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 US-08-669-536-1

Query Match 12.1%; Score 222.6; DB 2; Length 1812;
 Best Local Similarity 56.4%; Pred. No. 1.1e-48;

Matches 499; Conservative 0; Mismatches 374; Indels 12; Gaps 4;

972 AGATATCAAAAGCTGTTAAGAGATCCGAGCATTCGTGAGAGAGCCGTGCTGTT 1031
 634 AATATACCAAGATGAGAGAAAGACTCGTGGTCTTCGTGAGAACTACACATGAGTT 693
 1032 CACAGGCTGATGAGCTCTCAAGCTGGAGAGCTGTGTCTCGGCGCATGAGTCAAT 1091
 694 AAGAGGCTTATGATGAGAGAGCTAAGAACTTCTTCTCCGTATTAATGTTAT 753
 1092 GATTCTGTACCAAGAGATTTGATGATGATGATGATGATGATGATGATGATGAT 1151
 754 GATTCTGTACCAAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGAT 813
 1152 GGCCTGAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1211
 814 GGTCTATGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 873
 1212 GGTGAGTGAAGAGAGGCTGCTGCTGCTCTCAAGCTCTTGAGCAATTTGTAATT 1271
 874 GGAATGTCGCAAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 933
 1272 ACCGAATGCAACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1331
 934 ACCGAGTTGACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 993
 1332 CTAATGATGATCCGCAAGCTGATGATGATGATGATGATGATGATGATGATGAT 1391
 994 CTGAGAGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1053
 1392 GTGACACGAGAGCACTGTTGATGATGATGATGATGATGATGATGATGATGAT 1451
 1054 ATCATGCTTACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1113
 1452 TCCACACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1508
 1114 TTGTCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173
 1509 GTACGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1565
 1174 ATTAAGCTTAACACGACAGATGATGATGATGATGATGATGATGATGATGATGAT 1233
 1566 GCAAGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1622
 1234 GCTAGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1293
 1523 ATCAGAGCCACAAAGAGGCTTGTGCACTGATGATGATGATGATGATGATGATGAT 1679

DB 1294 TGCTCGTTCACTAACCAAGTCATTCGCCAAGCTGAGTTGTGGAATGAAGAAGAGAGTGGG 1353
 QY 1680 CGATACAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1739
 DB 1354 AAGATATGAGAAAGATGATGCTTGTGCAAAACCTCGACGGAAGGTTGCTGCACTT 1413
 QY 1740 CATCTGCATCATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799
 DB 1414 CATCTGCAGAACTCGAGCCAAAGCTTACCAAACTTTCGAAAGATCAAGCTGACTACATT 1473
 QY 1800 GGACTCAACAAATGGGCCATTCAAACTTAATTAATTAACAGATAC 1844
 DB 1474 AGCTTCAGATTGAGGCTTCCTTACAGCCCTGCTACACAGGTAC 1518

RESULT 6
 US-08-930-894-1
 Sequence 1, Application US/08930894
 Patent No. 6037524

GENERAL INFORMATION:
 APPLICANT: GREENLAND, Andrew James
 APPLICANT: DRAVER, John
 APPLICANT: SKIPSEY, Marc
 APPLICANT: WARNER, Simon
 TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEIN HYDROLASE PROMOTER
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pillsbury Madison & Sutro
 STREET: 1100 New York Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005-3918

COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: MS Word
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/930,894
 FILING DATE: 09-OCT-1997
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/GB96/00882
 FILING DATE: 10-APR-1996

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9507381.3
 FILING DATE: 10-APR-1995
 INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1767 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: SHH GENE FROM ASPARAGUS
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 26..1483
 OTHER INFORMATION: /codon_start= 26
 US-08-930-894-1

Query Match 11.6%; Score 214.6; DB 3; Length 1767;
 Best Local Similarity 55.8%; Pred. No. 1.3e-46;
 Matches 494; Conservative 0; Mismatches 379; Indels 12; Gaps 4;

972 AAGTATCAAAAGCTGTTAAGAGATCCGAGCATTTGTGAGAGAGAGGCTGCTGTT 1031
 596 AAGTACAGAGAGATGAGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655

QY	1032	CACAGCGTATAGCTCTCCAAAGCTGGGAAGCTGTGGTTCGGCCATGAAAGCTCAT	1091
Db	656	AAGAGCTTTACCAAGATGACGGCTAACAAFTTCCCTCTTTCCCTGCATCAATGTCAAT	715
QY	1092	GATTCTGTACCAACAGAAAGTTTGATTAACCTGTACTGCTGCCAGAAATCATTTTGAT	1155
Db	716	GACTCCGTCACAAAGACAGATTGTGACAAATCTGTATGATGACCGGACACTCTTCCCAT	775
QY	1152	GGCTGAAAGAGACACACAGTTGTGATGTTTGGTGGGAAACAAAGTGGTGTGTGGCTAT	1211
Db	776	GGTGTGTAGAGGGCCACTGATGTTATGATTTGGCTGGGAAGGTTGCACATTTCTGCGGAT	835
QY	1212	GGTAGGATAGGCAAGGGCTCTGTGCTGCTCTCAAAAGCTCTTGAGACATTTCTACATT	1277
Db	836	GGTATTTCCGAGAGAGGCTGTGCTGCTGCACACTAACACAGGCTGTGCCCTGTATTTGG	895
QY	1272	ACCGAAATGACACCCCATCTGTGCTGTGCACAGGCTGCATGATGGTGTTCACGGTGTAA	1331
Db	896	ACGGAGATTCACACCCCATCTGTGCTCTCAAGCCCTAATGGAGGGCTTTCAGGCTCTCAC	955
QY	1332	CTAATTAAGATCAATCCGGCAAGTGCAGATGTCATTAATCACTGGACACAGAAATTAAGATGA	1391
Db	956	CTCGAGATGTGTCTCAGAGGGGATATCTTTGTTACACACCGGTACAGAGACATC	1015
QY	1392	GTCACACGGGAGCCTTGATTCGCATGAAACAACTGTATGCTATGCAATATGGCCAC	1451
Db	1016	ATCATGCTGGACACAGAGAAAGATGAAGAAACAAATGCCATTGTCTGCAACATTTGGTAC	1077
QY	1452	TCCAAACACAGAAATCATGTGTGACACAGCTCCGAC--TCCGAGCTGACGTGGAGGGA	1507
Db	1076	TTTACACACAGATGTGACATGCTAGTTGGAGACATACCCTGGCATCAAGAGAAATCAC	1133
QY	1509	GTAGCTTCTGAGGTGAGACATATGTCATCTGGCCAGATGGCAA--ACGAGTGTGCTCCTG	1565
Db	1136	ATCAAGCCCGAGACTGACCGGTGGGTCTTCCCTGAAACCAACCTGTATTAATTTGTTCT	1199
QY	1566	GCAGAGGCTGCTACTCAATTTGAGCTGCTCCAGTT--CCACCTTGTGTTCTGCC	1622
Db	1196	GCTGAGGCGGAGCTCATGAACCTTGGGTGGTGCACATGCTACCCAGCTTTGTCATGTCC	1255
QY	1623	ATCACACCCACACACAGGCTTTGGCACATATGAACCTCTATAATG--CACCCAGGGG	1677
Db	1256	TGCTCTTTCACCAACCGAGGATTTGCTACACTGAGTTGGGATATAGAAGGACAGCGGC	1311
QY	1680	CGATACAGAGAGATGTGTACTTCTCTCTCTCAAAAAATGGATGAATACCTGTGCACCTG	1733
Db	1316	AAGTATGAGAAAGATTTACGTGTCTCCCAACACATCTGTATGATAGAAATAGAGGCGTT	1377
QY	1740	CATGTGCCATCATTTGATGCCACACTTACAGAGCTGACAGATGACCAAGCAAAATATCTG	1799
Db	1376	CACTTGGGCAAGCTCGGAGGCAAGCTTACAAACCTCAGCCCTTCACAGCGGGGACTCATC	1433
QY	1800	GGAATCAACAAAAATGGGCAATTCAAACCAATTAATTAAGATAC	1844
Db	1436	AGGCTCCCATCGAGGCTCCCTTACAGCCACTCATACAGGTAC	1480

RESULT 7
US-08-896-005-2

Sequence 2, Application US/08896005
Patent No. 5854023
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Puvi
TITLE OF INVENTION: S¹-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3114 Porter Drive
CITY: Palo Alto
STATE: CA

COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,005
FILING DATE: Filed Herewith
Prior APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0337 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2226 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADTUT04
CLONE: 1519044
US-08-896-005-2

	Query Match	6.3%	Score 117;	DB 2:	Length 2226;	
	Best Local Similarity	100.0%	Pred. No. 4.8e-21;			
	Matches 117;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1728 GTTGGCAGCTTCGATCGCCATCATTTGATGCCACCTTAGAGAGCTGCACATGACC	A	A	A	A	1787
Db	1 GTTGCGACTTCGCACTCGCCATCATTTGATGCCACCTTAGAGAGCTGCACATGACC	A	A	A	A	60
QY	1788 GCAAAATATCTGGGAGCTCAACAAAANNTGGGCACTCAAACCTTAATTATATACAGATAC	A	A	A	A	1844
Db	61 GC AAAATATCTGGGAGCTCAACAAAATGGGCACTCAAACCTTAATTATATACAGATAC	A	A	A	A	117

US-09-347-878-3
RESULT 8
Sequence 3, Application US/09347878C
Patent No. 6376210
GENERAL INFORMATION:
APPLICANT: Yuan, Chong
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
FILE REFERENCE: 25885-1651
CURRENT APPLICATION NUMBER: US/09/347,878C
CURRENT FILING DATE: 1999-07-06
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 2226
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(2226)
OTHER INFORMATION: Polynucleotide encoding human
OTHER INFORMATION: S-adenosyl-5-homocysteine hydrolase (SAHH) derived
OTHER INFORMATION: from bladder; n-a, c, g, or t
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 08/896,005
PATENT FILING DATE: 1997-07-17
PUBLICATION DATE: 1998-12-29
US-09-347-878-3

Query Match	6.38; Score 117; DB 4; Length 2226;
-------------	-------------------------------------

RESULT 11
US-09-081-395-8
Sequence 8, Application US/09081395
Patent No. 6083746
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive, 32nd Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,395
FILING DATE: 18-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 6083746nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-KK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-081-395-8
Query Match 3.9%; Score 72.6; DB 3; Length 289;
Best Local Similarity 58.6%; Pred. No. 7.9e-10;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 580 ACATCAAGCAGGAGATTTGGACGCGGAGATTGAGATTGAGAGACAGACATGCTG 639
DB 75 ACATCGGCGCTGCTGGCGGAGCAAGCCCTGGACATTGCTGAGAGAGATGCCGG 134
QY 640 CTGTGATTTCATCAGAGAAAGTGTCTCAGGGGAGAAAGCCCTTGCTGTCTAAATAG 699
DB 135 GCGTGTGCTGTATGCGGAGCGGTACTCGGCTCCAGCCAGCAAGGGGCGCCGATCG 194
QY 700 TGGGCTGTACACATCAGACCCAGACAGCGGTGTGATTGAGACACTGTGCTCCCTGG 759
DB 195 CTGGCTGCTGACATGACCGTGTGAGAGCGGCCGCTCATTTAGAGACCCCTGACCTGG 254
QY 760 GGGCTAGTCCGCTGTGCTGTCTGTTAATCTAC 794
DB 255 GTGCTGAGGTGCACTGTCTCAAGCTGCAACATCTTC 289

RESULT 12
US-09-416-833-8
Sequence 8, Application US/09416833
Patent No. 6197521
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Alligretti & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/416,833
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/204,740
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 6197521nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-416-833-8
Query Match 3.9%; Score 72.6; DB 4; Length 289;
Best Local Similarity 58.6%; Pred. No. 7.9e-10;
Matches 126; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 580 ACATCAAGCAGGAGATTTGGACGCGGAGATTGAGATTGAGAGACAGACATGCTG 639
DB 75 ACATCGGCGCTGCTGGCGGAGCAAGCCCTGGACATTGCTGAGAGAGATGCCGG 134
QY 640 CTGTGATTTCATCAGAGAAAGTGTCTCAGGGGAGAAAGCCCTTGCTGTCTAAATAG 699
DB 135 GCGTGTGCTGTATGCGGAGCGGTACTCGGCTCCAGCCAGCAAGGGGCGCCGATCG 194
QY 700 TGGGCTGTACACATCAGACCCAGACAGCGGTGTGATTGAGACACTGTGCTCCCTGG 759
DB 195 CTGGCTGCTGACATGACCGTGTGAGAGCGGCCGCTCATTTAGAGACCCCTGACCTGG 254
QY 760 GGGCTAGTCCGCTGTGCTGTCTGTTAATCTAC 794
DB 255 GTGCTGAGGTGCACTGTCTCAAGCTGCAACATCTTC 289

RESULT 13
PCT-US95-02521-8
Sequence 8, Application PC/TUS9502521
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-081-167A-6

Query Match 3.7%; Score 68.4; DB 3; Length 285;
Best Local Similarity 56.8%; Pred. No. 9.9e-09;
Matches 126; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

OY 573 GTGAAGAACATCAGCAGGAGCAATTTGGACGCCGGAGATTGAGATTGCAGAGCAAGAC 632
DB 64 GTCCGGGACATGAGACTGGCCCGCTGGGGAGGAGGCTCTGATATGACTGAGAAATGAG 123
OY 633 ATGTCTGCTGTATTTCACTCAGAAACGCTGCTCAGGGGGAGAGCCCTTGGCTGCTGCT 692
DB 124 ATGCCAGGGTTGATCGCGCATGCGGAGATGTAATCACTCAGCCCTCCAGGCACTGAAAGGCT 183
OY 693 AAAATAGTGGGCTGTACACATCAGACAGCCAGAGCGGTGTGATTGAGACACTCTGT 752
DB 184 CGCATTTGCTGGCTGCTGCGCATGACCGTGGAGACTGCTGTCTCATTTGAGACTCTCGTG 243
OY 753 GCCCTGGGGGCTCAGTGCCTGCTGCTGCTGTAACATCTAC 794
DB 244 GCCCTGGGCTGAGAGCGGGGTGTCACAGCTGCAACATCTTC 285

Search completed: April 21, 2003, 05:16:14
Job time : 2569.43 secs


```

Db 121 CGAGCTCAGCAGGAGGCGGCGGAGAGGCTGGGCGATCGCGTGTGAGAGGCGCGCGC 180
QY 181 GGGCAGGCGGGGCGGCGGCGGAGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db 181 GGGCAGGCGGGGCGGCGGCGGAGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 241 GGGCGGGGCGGAGAGTGCATGCTGACCGGATGCGCGCTGCGCGGCGGCGGCGGCGGCGG 300
Db 241 GGGCGGGGCGGAGAGTGCATGCTGACCGGATGCGCGCTGCGCGGCGGCGGCGGCGGCGG 300
QY 301 TGAAGCAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 301 TGAAGCAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 AGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 AGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 CCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 421 CCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 GTTACAGTGCATCTTACACAGATAGCTGATGATGATGATGATGATGATGATGATGATGAT 540
Db 481 GTTACAGTGCATCTTACACAGATAGCTGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 AAACCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 AAACCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 GAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 GAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 GTGCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 GTGCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 CCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 CCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 CTGTGACATCTTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 840
Db 781 CTGTGACATCTTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 840
QY 841 CAGTGTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 CAGTGTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 TGAACATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 901 TGAACATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 GGGTGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 961 GGGTGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 TGAAGTGTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TGAAGTGTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 TGAAGTGTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 TGAAGTGTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 CCATTTTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1141 CCATTTTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 TGTGTGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260

```

```

Db 1201. TGTGTGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 TGTGTGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 TGTGTGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 GGGTGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 GGGTGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 ATTAAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 ATTAAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 ATATGGGCGACCTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1441 ATATGGGCGACCTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 GGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 1501 GGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1561 TCCGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 TCCGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 CCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1621 CCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1681 GATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1681 GATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
QY 1741 ATCTGCGCATATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 1741 ATCTGCGCATATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 1801 GACTCAACAAATAATGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1844
Db 1801 GACTCAACAAATAATGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1844

```

```

RESULT 2
US-09-925-300-439
; Sequence 439, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 439
; LENGTH: 721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (688)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-439

```

```

Query Match 38.1%; Score 702.4; DB 10; Length 721;
Best Local Similarity 98.0%; Pctid: No. 2, 9e-209;
Matches 703; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

```

```
QY 1 GGGCGGGGAGGTCGAGAGCTCGAGAGCTGCTTGTGTTCTGTGTGTGCGCCAGCGTGCCT 60
    |||||
DB 4 GGGCGGGGAGGTCGAGAGCTCGAGAGCTGCTTGTGTTCTGTGTGTGCGCCAGCGTGCCT 63
QY 61 GTCCCGCTGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
    |||||
DB 64 GTCCCGCTGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 123
QY 121 CCACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
    |||||
DB 124 CCACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 183
QY 181 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
    |||||
DB 184 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 243
QY 241 GGGCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
    |||||
DB 244 GGGCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 303
QY 301 TGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
    |||||
DB 304 TGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
QY 361 AGCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
    |||||
DB 364 AGCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
QY 421 CCAAAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
    |||||
DB 424 CCAAAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 483
QY 481 GTTCAAGCTGATCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
    |||||
DB 484 GTTCAAGCTGATCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 543
QY 541 AAACCAAGCTGATCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
    |||||
DB 544 AAACCAAGCTGATCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 603
QY 601 GAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
    |||||
DB 604 GAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663
QY 661 GTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717
    |||||
DB 664 GTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
```

RESULT 3

```
US-09-867-701-2023
; Sequence 2023, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Agate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2023
; LENGTH: 553
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(553)
; OTHER INFORMATION: n - A,T,C or G
```

US-09-867-701-2023

```
Query Match 25.0%; Score 461.4; DB 10; Length 553;
Best Local Similarity 97.3%; Pred. No. 6, 2e-134;
Matches 510; Conservative 0; Mismatches 9; Indels 5; Gaps 4;
```

```
QY 1322 GGTGTAAAGCTAAATGAATGATCCGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1381
    |||||
DB 1 GGTGTAAAGCTAAATGAATGATCCGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60
QY 1382 TTAGATGTAGTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1441
    |||||
DB 61 TTAGATGTAGTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 1442 TATGGGCACTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501
    |||||
DB 121 TATGGGCACTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 1502 GGAAGCAGTACGTTCTCAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1561
    |||||
DB 181 GGAAGCAGTACGTTCTCAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 239
QY 1562 CCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
    |||||
DB 240 CCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 299
QY 1622 CATCAGAGCAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
    |||||
DB 300 CATCAGAGCAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
QY 1682 ATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
    |||||
DB 360 ATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
QY 1741 ATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
    |||||
DB 420 ATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478
QY 1801 GACTCAACAAAATTTGGGCACTTCAACCTAATTTATTTATTTATTTATTTATTTATTT 1844
    |||||
DB 479 GACTCAACAAAATTTGGGCACTTCAACCTAATTTATTTATTTATTTATTTATTTATTT 520
```

RESULT 4

```
US-09-879-536-850/c
; Sequence 850, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 850
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(636)
OTHER INFORMATION: n = A,T,C or G
US-09-879-536-850

Query Match 22.0%; Score 406.4; DB 10; Length 636;
Best Local Similarity 99.5%; Pred. No. 1.1e-116;
Matches 418; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 707 TACACATACAGCCAGACAGC-GGTGTGATTGAGACACTGTGCTGGGGCTC 765
DB 420 TACACATACAGCCAGACAGCGGGTGTGATTGAGACACTGTGCTGGGGCTC 361
QY 766 AGTCCGCTGTCTGCTTTGATACATCTACTCACTAGAAAGTGTGACACTGG 825
DB 360 AGTCCGCTGTCTGCTTTGATACATCTACTCACTAGAAAGTGTGACACTGG 301
QY 826 CTGAGGCTGAGTGTGAGTGTGCTGGAAGGCGAGTCAAGATGACTTGTGCT 885
DB 300 CTGAGGCTGAGTGTGAGTGTGCTGGAAGGCGAGTCAAGATGACTTGTGCT 241
QY 886 GTATTGACCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 945
DB 240 GTATTGACCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 181
QY 946 GAGACTTACCCACTGGGTTTATAGAGTATCCAAAGTGTGTTAGAGATCCGAGCA 1005
DB 180 GAGACTTACCCACTGGGTTTATAGAGTATCCAAAGTGTGTTAGAGATCCGAGCA 121
QY 1006 TTGGGAGAGAGAGGCTGATGCTGATGATGATGATGATGATGATGATGATG 1065
DB 120 TTGGGAGAGAGAGGCTGATGCTGATGATGATGATGATGATGATGATGATG 61
QY 1066 TCTGTGTTCCGCGCATGAGTCAATGATGATGATGATGATGATGATGATG 1125
DB 60 TCTGTGTTCCGCGCATGAGTCAATGATGATGATGATGATGATGATGATGATG 1

RESULT 5
US-09-925-301-217
Sequence 217, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 217
LENGTH: 2200
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1218)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-217

Query Match 21.8%; Score 402.4; DB 10; Length 2200;
Best Local Similarity 58.4%; Pred. No. 3.8e-115;
Matches 741; Conservative 0; Mismatches 521; Indels 6; Gaps 2;

QY 580 ACATCAAGCAGCAGATTTGAGAGCGGAGATGATGATGAGAGCAGATGCTG 639
DB 108 ACATCGCCCTGCTGCTGCTGAGAGCGAGCGCTGAGCATTTGCTGAGAGAGATGCGG 167

QY 640 CTCGTATTTCACTAGAGAAAGCTCTCAGGGGAGAAAGCCCTTGGCTGCTAAATAG 699
DB 168 GCTGTATGCGATATGCGAGAGGAGTACTCGGCTCCAAAGCCACTAGAGGCGCGCATCG 227
QY 700 TGGCTGTACACATACAGACAGCCAGAGAGGAGTGTGATGATGAGACACTGTGCGCGG 759
DB 228 CTGCTCTCTCAGATATGACCTGAGAGAGGCGCTCTCTCTATTTGAAACCTCTCACCTTGG 287
QY 760 GGGCTCAGTGGCGGTGTGCTGCTTTGATACATCTACTCACTAGAAAGTGTGAG 819
DB 288 GTGTGAGGTGCAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 347
QY 820 CACTGCTGAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 879
DB 348 CCATTTGCAAGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 407
QY 880 GGTGTGTATGAGCAGCTGTGATGATGATGATGATGATGATGATGATGATGATG 939
DB 408 TGTGTGTATGAGCAGCTGTGATGATGATGATGATGATGATGATGATGATGATG 467
QY 940 ATGGGAGAGACTTAAACCACTGGGTTTATAGAGTATCCAAAGTGTGTTAGAGATCC 999
DB 468 ACGGGGCGAGCTCAGCAACCTCATCCAGACAGTATCCGAGCTTCTGCGAGCATCC 527
QY 1000 GAGCATTTGAGAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATG 1059
DB 528 GAGCATTTCTGAGAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATG 587
QY 1060 GAAAGCTGTGTGCTGCGGCGCATGAGTGTGATGATGATGATGATGATGATGATG 1119
DB 588 GATATCTTCAAGTGTGCTGCGGCGCATGAGTGTGATGATGATGATGATGATGATGATG 647
QY 1120 ACTGTACTGTGCGGCGCATGAGTGTGATGATGATGATGATGATGATGATGATGATG 1179
DB 648 ACTGTACTGTGCGGCGCATGAGTGTGATGATGATGATGATGATGATGATGATGATG 707
QY 1180 TTGTGAGAGAAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1239
DB 708 TTGTGAGAGAAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 767
QY 1240 CTCATCAAGCTCTGAGAGCATATGCTACATTTACGAAATGAGCCCATGCTGTCTGTC 1299
DB 768 CCTGTGCGGCTTGTGAGAGCGCGCTCATCATCCGAGATGAGCCCATGAGCCACTGTC 827
QY 1300 AGGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
DB 828 AGGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 887
QY 1360 TCGTATTAATCTGACAGGAAATAGATGATGATGATGATGATGATGATGATGATG 1419
DB 888 TCGTATTAATCTGACAGGAAATAGATGATGATGATGATGATGATGATGATGATG 947
QY 1420 AAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1479
DB 948 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1007
QY 1480 TCCGCACTCCGAGAGTACAGTGTGAGAGAGAGTGTGATGATGATGATGATGATG 1539
DB 1008 TCAACAGAGAGAGAGTGTGAGAGAGAGTGTGATGATGATGATGATGATGATG 1067
QY 1540 CAGATGCAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1599
DB 1068 AGATGAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1127
QY 1600 CAG---TTCCACCTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1656
DB 1128 TGGGCAACCCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1187
QY 1657 AACTCTTAATGACAGCGAGAGAGAGTGTGATGATGATGATGATGATGATGATGATG 1716
DB 1188 AGCTGTGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1244
QY 1717 TGTATGATACGTTTGCAGCTTGTGATGATGATGATGATGATGATGATGATGATG 1776

Db 1245 TGGATGAGCAGTGGCTGAAGCCACCCTGGCAGAGCTGAATGTGAATTGACCAAGCTAA 1304
 QY 1777 CAGATGACCAAGCAAAATATCTGGAGCTCAACAAAATGGGCACTTCAAACTTAATATT 1836
 Db 1305 CTGAGAGCAAGCAAGCCAGTACTGCGCATCTCTGTATGAGCCCTTCAAGCCGATACT 1364
 QY 1837 ACAGATAC 1844
 Db 1365 ACCGCTAC 1372

RESULT 6
 US-10-044-090-344
 ; Sequence 344, Application US/10044090
 ; Patent No. US20020137081A1

; GENERAL INFORMATION:
 ; APPLICANT: Oiga Bandman
 ; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
 ; FILE REFERENCE: PA-0028 US
 ; CURRENT APPLICATION NUMBER: US/10/044, 090
 ; CURRENT FILING DATE: 2002-01-09
 ; NUMBER OF SEQ ID NOS: 850
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 344
 ; LENGTH: 2429
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; OTHER INFORMATION: Incyte ID No. US20020137081A1 1468237CBI
 US-10-044-090-344

Query Match 21.8%; Score 402.4; DB 12; Length 2429;
 Best Local Similarity 58.4%; Pred. No. 4e-115;
 Matches 711; Conservative 0; Mismatches 521; Indels 6; Gaps 2;

QY 580 ACATCAGCAGGAGATTTGAGCCGCGGAGATTGAGATTGACAGCAAGACATGCTG 639
 Db 108 ACATCGCCGCTGCTGCTGGGAGCAAGGCCCTGCAATGCTGAGAGAGATGGCGG 167
 QY 640 CTGTGATTTCACTGAGAAACGTGCTGAGGGGAGAAAGCCCTTGGCTGCTTAAATAG 699
 Db 168 GCGTGTGCTGATGCGGAGCGGTACTGCGCTCCAAAGCCACTGAAAGGGCCCGCATCG 227
 QY 700 TGGGCTGATACACATACAGCCAGCCAGAGCGGCTGATGATGAGACACTGCTGCCCTGG 759
 Db 228 CTGGCTGCTGACATGACCGTGGAGAGCGCGCTCTCATTTGAGACCTCTGCAACCTGG 287
 QY 760 GGGCTGAGTCCGCTGCTGCTGTTGTAACATCTACTCAACTCAGATGAATGATGCTGAG 819
 Db 288 GTGCTGAGGTGCAAGTGTCCAGCTGCAACATCTTCTCCACCCAGACCATGCGGCGCTG 347
 QY 820 CACTGCTGAGGCTGAGATTGCAAGTGTCCCTTGGAAAGGCGAGTCAAGATGACTTCT 879
 Db 348 CCAATGCGCAAGGCTGCGATTCGGGTATGCTGCTGAAAGGGGAAAGGAGACAGATGACC 407
 QY 880 GGTGTGATTTGACCGCTGTGTGTAACATGATGAGTGGGTGGCAGCCCAATGATCTGATG 939
 Db 408 TGTGTGCTCATTTGAGCAGACCTGTACTTCAAGAGGGGCCCTTAACATGATTTGGACG 467
 QY 940 ATGAGGAGAGACTTAACCACTGGGTTTATTAAGATATCCAAAGCTGTTTAAAGATCC 999
 Db 468 ACGGGGGGCGACTCACCACCTCTCATCCACCAAGTACCCGACACTTCTGCAAGCATCC 527
 QY 1000 GAGGCTATTGAGAGAGAGCGGTGACTGTGTACAGGCTGTATCAGCTTCCAAAGCTG 1059
 Db 528 GAGGCAATCTGTGAGAGAGACAGACGCTGGGGTCCACAACCTCTACAAGATGATGGCAATG 587
 QY 1060 GGAAGCTGTGTTCGGGCACTGAAGCTCAATGATCTGTGTTACCAAGAGATTGATA 1119
 Db 588 GGATCTCAAGGTGCTCCGACATCAATGTCAATGACTCCGTGACCAAGAGAGATTGACA 647

QY 1120 ACTTGTACTGCTGCCGGAATTCATTTTGGATGGCCCTGAAGAGACCAAGATGATGCT 1179
 Db 648 ACCCTATGCTGCTGCCGAGAGTCCCTCATAGATGATCAAGCGGCCACAGATGTGATGA 707
 QY 1180 TTGGTGGAAACAAGTGTGTGTGTGTGCTATGAGTGTGAGTGTGAGAGAGGCTGTGCTG 1239
 Db 708 TTGCGGCAAGGTATGCGGT 767
 QY 1240 CTCTCAAGGCTCTTGGAGCAATTTGTCAATTTACCGAATTCAGCCCATCTGTCTGTC 1299
 Db 768 CCGTGGGGGGTTTGGAGCCCGGCTCATTCATTCACAGATGTGACCCATCAAGCATGTC 827
 QY 1300 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
 Db 828 AGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 887
 QY 1360 TCGTAAATGCTGCAAGAGAAATGAATGATGATGATGATGATGATGATGATGATGATGAT 1419
 Db 888 TCTTTGTCACCAACAGGCTGT 947
 QY 1420 AAACAGT 1479
 Db 948 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1007
 QY 1480 TCCGCACTCCGAGCTGACGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1539
 Db 1008 TCAACGAGAACGCGGTGAG 1067
 QY 1540 CAGATGCAAG 1599
 Db 1068 AGATGAG 1127
 QY 1600 CAG---TTCCACCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1656
 Db 1128 TGGGCAACCCAGGCTGT 1187
 QY 1657 AACTCTATATGACCCGAGGCGAGATGCAAGAGATGTACTTGTCTTAAAGAAA 1716
 Db 1188 AGCTGTGAGCCCATTCAGACAG---TACCCGTTGGGGTTCATTTCTGCGCCCAAGAGC 1244
 QY 1717 TGGATGATATGCTGACAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1776
 Db 1245 TGGATGAG 1304
 QY 1777 CAGATGACCAAGCAAAATATCTGGAGCTCAACAAAATGGGCACTTCAAACTTAATATT 1836
 Db 1305 CTGAGAGCAAGCCCAAGTACTGCGCATCTCTGTGATGAGCCCTTCAAGCCGATGACT 1364
 QY 1837 ACAGATAC 1844
 Db 1365 ACCGCTAC 1372

RESULT 7
 US-09-880-107-346
 ; Sequence 346, Application US/09880107
 ; Patent No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 ; FILE REFERENCE: 44921-5028-WO
 ; CURRENT APPLICATION NUMBER: US/09/880, 107
 ; CURRENT FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211, 379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237, 054
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: Patent Ver. 2.1
 ; SEQ ID NO 346

QY	1431	ATCTATGCAATATATGGCCACTCCAAACAGAAATGATGTG---ACAGCCTCCGCACT	1487
Db	1006	CTGTGGGCAACATCCGGTCACTTTGATTAATGAATGATATCATCTCCCTGTTGGACGC	1065
QY	1488	CCGAGGCGAGAGTGGGAGCAGTAGCTTTCAGGTGGACCATGTATCTGGCCAGATGGC	1547
Db	1066	GACACGTCACCCGCACACAGATCAAGCCACAGAGTGCAGAGTTCACTTCTCCACGGT	1125
QY	1548	AAACGAGTTGTCTCTCTGGCAGAGGGTGTCTACTCAATTGAGCGTCCAC---AGTT	1604
Db	1126	CGCTGCATCATCGTCGTCTCGAAGGCGCGCTTTAACTTTGGCAAGCCACCGAGAC	1185
QY	1605	CCCAACCTTTGTCTGTCCATCAGACCACAAACAGAGCTTTGGCACTGATAGACTCTAT	1666
Db	1186	CCATCATTTGTCATGTCCTCAACTCTTTGCGCGATCAACACATTCGCGAGATCCAACTTTC	1245
QY	1665	AATGCACCCGAGGGGCGCATACAGACAGATGTGTACTTGTCTCTTAAGAAATGATGAA	1722
Db	1246	CA---AAACGAGAGCAGTACGAGAAAGAGAGTGTACCGTGTGCTTAAGGTTCTCGCGAA	1302
QY	1725	TACGTGCACACTGTCATCTGCGCATCAATTTGATAGGCCCACTTCACAGAGTGACAGATGAC	1788
Db	1303	AAGGCGACCATTCACAGTTTGAAGGCTCTTGCGGGGTAGCTACACGAACTGCACAAAGAG	1363

DD 938 CAGGCGCICGGLCAGGICACGAA

Db 938 CAGGGCGCTGCGGCTCAAGGTCACCGAAGCTGACCCCATCATCAAGGCTCTTCAGGCTCTGATG 997

QY 1311 GATGGTTAGGGTGTAAAGCTAAATGAATCCGGCAAGTCGATGCTAAATACT 1370
 DB 998 GATGGCTACTGCTGTGACCGTTGATGAGCCATGAGACCGGCAGCATCGATGATACC 1057
 QY 1371 TGCACAGGAATAGATAGTAGACACGGGACACTTGGATCGCATGAAAACAGTTGT 1430
 DB 1058 GCGACGGGCAACAGACATCATTTCTTCTGACAGATGCTCAAGATGAGATGACGCT 1117
 QY 1431 ATGTAATGCAATATGGGCACTCCACACAGAAATGATGATG--ACACGCTCCGCACT 1487
 DB 1118 CTGCTGGGCAACATCGGTCTTGTGATTAAGATGATGATGCTTCTGTTGACCGC 1177
 QY 1488 CCGGACCTGACGTGGAGCGAGTACGTTCTCAGGTGAGACCATGATCTGCGCAGATGCG 1547
 DB 1178 GACGACGTCACCGCACACAGATCAAGCCACAGAGTGCAGAGTTCACTTCACCGGCT 1237
 QY 1548 AAGGAGTTCTCTCTGCGAGAGGCTGCTACTCAATTTGAGCTGCTCAG--AGTT 1604
 DB 1238 CGCTCATCATCTGCTGCTGCGAAGGTCGCTGTTGAACCTTGGCAACGCGCAGCGAC 1297
 QY 1605 CCCACCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1664
 DB 1298 CCATCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1357
 QY 1665 AATGACCCGAGGGGCGATCAACAGAGATGTGTAATGCTGCTGCTGCTGCTGCTGCTGCT 1724
 DB 1358 CA---AAGCGAAGAGAGTACGAGAGAGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCT 1414
 QY 1725 TAGTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1784
 DB 1415 AAGGTGACGACGATCCAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1474
 QY 1785 CAGCAAAATATCTGGGACTCAACAAAATGGGCACTTCAACCTAATATATATACAGATAC 1844
 DB 1475 CAGGCTGAGTACATCGGCGTTGAGCTTGCAGGCCCATTCAGCGCGAGACACTACCGCTAC 1534

RESULT 11
 US-10-021-121-1/c
 ; Sequence 1, Application US/10021121
 ; Patent No. US7002014244A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Caras, Ingrid W
 ; TITLE OF INVENTION: A2-1 Neurotrophic Factor
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Genentech, Inc.
 ; STREET: 1 DNA Way
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 Inch, 1.44 Mb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Winpatin (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/021,121
 ; FILING DATE: 06-Dec-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/635,130
 ; FILING DATE: 19-Mar-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Torchia, Ph.D., Timothy E.
 ; REGISTRATION NUMBER: 36,700
 ; REFERENCE/DOCKET NUMBER: P1001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650/225-8674
 ; TELEFAX: 650/952-9681
 ; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1877 base pairs
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Double
 ; TOPOLOGY: Linear
 ; FEATURE:
 ; NAME/KEY: Extra Cellular Domain
 ; LOCATION: 244-899
 ; IDENTIFICATION METHOD:
 ; OTHER INFORMATION:
 ; FEATURE:
 ; NAME/KEY: Transmembrane Domain
 ; LOCATION: 901-978
 ; IDENTIFICATION METHOD:
 ; OTHER INFORMATION:
 ; FEATURE:
 ; NAME/KEY: signal peptide
 ; LOCATION: 244-321
 ; IDENTIFICATION METHOD:
 ; OTHER INFORMATION:
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-10-021-121-1
 Query Match 10.3%; Score 190; DB 12; Length 1877;
 Best Local Similarity 100.0%; Pred. No. 9,8e-49;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1283 CCCCATCTGCTGCTGAGGCGTCGATGATGGGTTCAGGGGTGTAAGCTAAATGAGT 1342
 DB 1446 CCCCATCTGCTGCTGAGGCGTCGATGATGGGTTCAGGGGTGTAAGCTAAATGAGT 1387
 QY 1343 CATCCGCAAGTGCATGCTGTAATTAATCTGACAGAGAAATAGATAGTACACCGGA 1402
 DB 1386 CATCCGCAAGTGCATGCTGTAATTAATCTGACAGAGAAATAGATAGTACACCGGA 1327
 QY 1403 GCACCTTGATGCGATGAAAACGTTGTATGCTATGCAATATGGGCCATCCACACAGA 1462
 DB 1326 GCACCTTGATGCGATGAAAACGTTGTATGCTATGCAATATGGGCCATCCACACAGA 1267
 QY 1463 AATCGATGTG 1472
 DB 1266 AATCGATGTG 1257

RESULT 12
 US-09-746-660A-101
 ; Sequence 101, Application US/09746660A
 ; Publication No. US20030049804A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompeius, Markus
 ; APPLICANT: Kroege, Burkhard
 ; APPLICANT: Schroder, Hartwig
 ; APPLICANT: Zelder, Oskar
 ; APPLICANT: Haberhauser, Gregor
 ; APPLICANT: Kim, Jun-Won
 ; APPLICANT: Lee, Heung-Schick
 ; APPLICANT: Hwang, Byung-Joon
 ; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
 ; FILE REFERENCE: BGI-121CP2
 ; CURRENT APPLICATION NUMBER: US/09/746,660A
 ; CURRENT FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 09/606740
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 09/603124
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 60/141031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: 60/142101
 ; PRIOR FILING DATE: 1999-07-02
 ; PRIOR APPLICATION NUMBER: 60/148613
 ; PRIOR FILING DATE: 1999-08-12
 ; PRIOR APPLICATION NUMBER: 60/187970

; CURRENT FILING DATE: 2002-01-04
 ; PRIOR APPLICATION NUMBER: 06/260,040
 ; PRIOR FILING DATE: 2001-01-05
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: Patent version 3.0
 ; SEQ ID NO 26
 ; LENGTH: 3830
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; US-10-037-598-26

Query Match 9.3%; Score 172.4; DB 9; Length 3830;
 Best Local Similarity 55.2%; Pred. No. 4,6e-43;
 Matches 423; Conservative 0; Mismatches 331; Indels 12; Gaps 4;

QY 1091 TGAATCTGTACCAACAGAGTTGATTAATCTGTCTGCTGCGGAGAAATCCATTGGA 1150
 DB 3062 TTAATCTGTGTTTGTATACAGTTGATGCAACTGTATGGGTGCCGTCCTCTCCCTGA 3121
 QY 1151 TGGCTGAAGAGGACACAGATGATGTTGGTGGGAAACAGTGGTGTGTGGCTA 1210
 DB 3122 TGGTCTCATGAGGGCTACCGATGTATGATGCTGGAAGAGTGGTGTGTGGCTGATTA 3181
 QY 1211 TGGTGAAGTGAAGGAGGCTGCTGCTGCTGCTCAAAAGCTCTTGGAGCAATTGTCAAT 1270
 DB 3182 TGGTGAAGTGAAGGAGGCTGCTGCTGCTGCTCAAAAGGAGGCTGCTGCTGCTCAATCGT 3241
 QY 1271 TACCGAATCGAACCCCATGTGCTGCTGCAAGGCTGATGATGGGTGAGGTGGTAA 1330
 DB 3242 GACCGAGATGATCCCATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAC 3301
 QY 1331 GCTAAATGAAGTATCCGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1390
 DB 3302 CTGGAGGATGTTGTTCTGAGGCTGATATCTTGTACACACACCGGTAACAAGACAT 3361
 QY 1391 AGTGAACGGGACCATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1450
 DB 3362 CATCATGTTGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1507
 QY 1451 CTCACACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1507
 DB 3422 CTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3481
 QY 1508 AGTACGTTCTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1564
 DB 3482 CATCAAGCCCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3541
 QY 1565 GGCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1621
 DB 3542 GGCCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3601
 QY 1622 CATCAACAGGACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1678
 DB 3602 CTGCT 3661
 QY 1679 GCGATCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1738
 DB 3662 CAAGTCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3721
 QY 1739 GGAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1798
 DB 3722 TCACTGAGGCAAACTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3781
 QY 1799 GGGACCAACAAATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1844
 DB 3782 CAGTGTGCTGTTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3827

RESULT 15
 US-09-754-853A-4/c
 ; Sequence 4, Application US/09754853A
 ; Publication No. US20030005491A1
 ; GENERAL INFORMATION:

; APPLICANT: Hauge, Brian M.
 ; APPLICANT: Parnell, Laurence D.
 ; APPLICANT: Parsons, Jeremy D.
 ; APPLICANT: Wang, Ming Li
 ; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
 ; FILE REFERENCE: 38-10(15810)B
 ; CURRENT APPLICATION NUMBER: US/09/754,853A
 ; PRIOR FILING DATE: 2001-01-05
 ; PRIOR APPLICATION NUMBER: US 60/174,880
 ; NUMBER OF SEQ ID NOS: 1119
 ; SEQ ID NO 4
 ; LENGTH: 513509
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (11805)..(11968)..(114684)..(115204)
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(513509)
 ; OTHER INFORMATION: unsure at all n locations
 ; OTHER INFORMATION: Clone ID: 318013-Region_A3
 ; US-09-754-853A-4

Query Match 9.3%; Score 172.4; DB 9; Length 513509;
 Best Local Similarity 55.2%; Pred. No. 6,6e-42;
 Matches 423; Conservative 0; Mismatches 331; Indels 12; Gaps 4;

QY 1091 TGAATCTGTACCAACAGAGTTGATTAATCTGTCTGCTGCGGAGAAATCCATTGGA 1150
 DB 178978 TTAATCTGTGTTTGTATACAGTTGATGCAACTGTATGGGTGCCGTCCTCTCCCTGA 178919
 QY 1151 TGGCTGAAGAGGACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1210
 DB 178918 TGGTCTCATGAGGGCTACCGATGTATGATGCTGGAAGAGTGGTGTGTGGCTGATTA 178959
 QY 1211 TGGTGAAGTGAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1270
 DB 178958 TGGTCTCATGAGGGCTACCGATGTATGATGCTGGAAGAGTGGTGTGTGGCTGATTA 178999
 QY 1271 TACCGAATCGAACCCCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1330
 DB 178998 GACCGAGATGATCCCATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 178739
 QY 1331 GCTAAATGAAGTATCCGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1390
 DB 178738 CTGGAGGATGTTGTTCTGAGGCTGATATCTTGTACACACACCGGTAACAAGACAT 178679
 QY 1391 AGTGAACGGGACCATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1450
 DB 178678 CATCATGTTGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 178619
 QY 1451 CTCACACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1507
 DB 178618 CTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 178559
 QY 1508 AGTACGTTCTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1564
 DB 178558 CATCAAGCCCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 178499
 QY 1565 GGCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1621
 DB 178498 GGCCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 178439
 QY 1622 CATCAACAGGACACAGGCTTGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1678
 DB 178438 CTGCTGCTTCAACACAGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 178379
 QY 1679 GGGATCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1738
 DB 178378 CAAGTCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 178319

